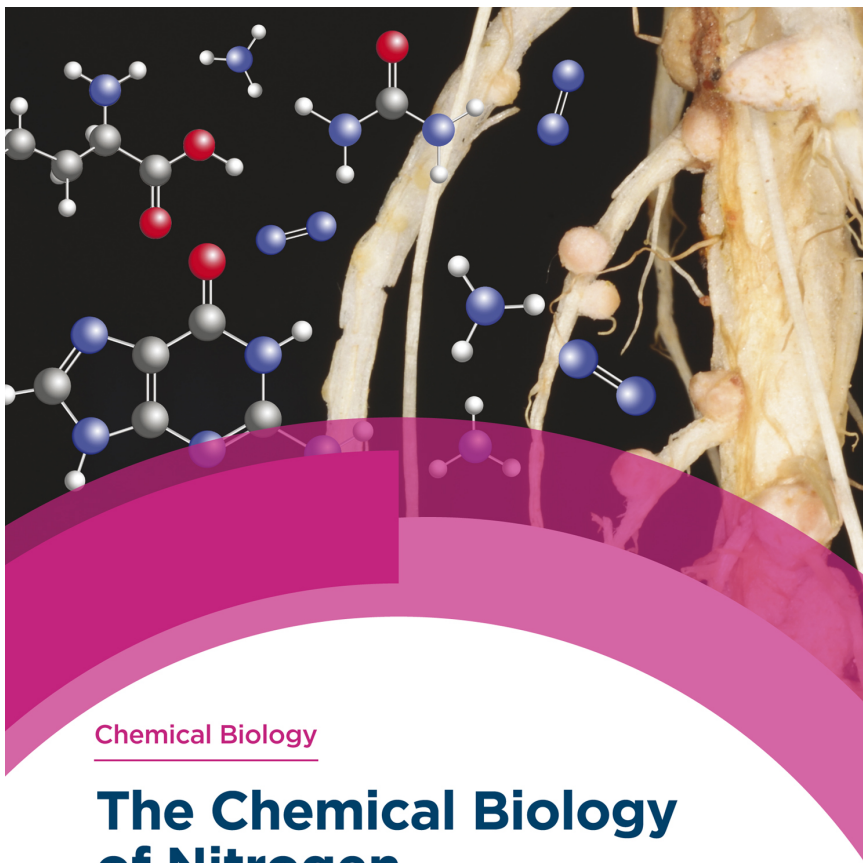




Chemical Biology

The Chemical Biology of Nitrogen

Christopher T. Walsh

The Chemical Biology of Nitrogen

Chemical Biology

Editor-in-chief:

Kira J. Weissman, *Lorraine University, France*

Titles in the series:

- 1: High Throughput Screening Methods: Evolution and Refinement
- 2: Chemical Biology of Glycoproteins
- 3: Computational Tools for Chemical Biology
- 4: Mass Spectrometry in Chemical Biology: Evolving Applications
- 5: Mechanisms of Primary Energy Transduction in Biology
- 6: Cyclic Peptides: From Bioorganic Synthesis to Applications
- 7: DNA-targeting Molecules as Therapeutic Agents
- 8: Protein Crystallography: Challenges and Practical Solutions
- 9: Oxidative Folding of Proteins: Principles, Biological Regulation and Design
- 10: Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology
- 11: Synthetic Glycomes
- 12: The Chemical Biology of Sulfur
- 13: The Chemical Biology of Phosphorus
- 14: DNA Damage, DNA Repair and Disease: Volume 1
- 15: DNA Damage, DNA Repair and Disease: Volume 2
- 16: The Discovery and Utility of Chemical Probes in Target Discovery
- 17: Inhibitors of Protein-Protein Interactions
- 18: The Chemical Biology of Nitrogen

How to obtain future titles on publication:

A standing order plan is available for this series. A standing order will bring delivery of each new volume immediately on publication.

For further information please contact:

Book Sales Department, Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, CB4 0WF, UK

Telephone: +44 (0)1223 420066, Fax: +44 (0)1223 420247

Email: booksales@rsc.org

Visit our website at www.rsc.org/books

The Chemical Biology of Nitrogen

By

Christopher T. Walsh

Stanford University, USA

E-mail: cwalsh2@stanford.edu



Print ISBN: 978-1-83916-392-0

PDF ISBN: 978-1-83916-433-0

EPUB ISBN: 978-1-83916-434-7

Print ISSN: 2055-1975

Electronic ISSN: 2055-1983

A catalogue record for this book is available from the British Library

© Christopher T. Walsh 2021

All rights reserved

Apart from fair dealing for the purposes of research for non-commercial purposes or for private study, criticism or review, as permitted under the Copyright, Designs and Patents Act 1988 and the Copyright and Related Rights Regulations 2003, this publication may not be reproduced, stored or transmitted, in any form or by any means, without the prior permission in writing of The Royal Society of Chemistry or the copyright owner, or in the case of reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency in the UK, or in accordance with the terms of the licences issued by the appropriate Reproduction Rights Organization outside the UK. Enquiries concerning reproduction outside the terms stated here should be sent to The Royal Society of Chemistry at the address printed on this page.

Whilst this material has been produced with all due care, The Royal Society of Chemistry cannot be held responsible or liable for its accuracy and completeness, nor for any consequences arising from any errors or the use of the information contained in this publication. The publication of advertisements does not constitute any endorsement by The Royal Society of Chemistry or Authors of any products advertised. The views and opinions advanced by contributors do not necessarily reflect those of The Royal Society of Chemistry which shall not be liable for any resulting loss or damage arising as a result of reliance upon this material.

The Royal Society of Chemistry is a charity, registered in England and Wales, Number 207890, and a company incorporated in England by Royal Charter (Registered No. RC000524), registered office: Burlington House, Piccadilly, London W1J 0BA, UK, Telephone: +44 (0) 207 4378 6556.

For further information see our web site at www.rsc.org

Printed in the United Kingdom by CPI Group (UK) Ltd, Croydon, CR0 4YY, UK

Preface

Nitrogen is omnipresent in our lives. Some 4000 trillion tons of nitrogen gas, N_2 , surround us as four-fifths of the global atmosphere, invisible, tasteless, inert. We breathe it in and out some eight thousand times a day without beneficial or harmful consequence (as long as we are at one atmosphere pressure). All this N_2 gas is unavailable to us. Yet, we humans have up to five kilograms of nitrogen atoms in our bodies, embedded in carbon-based organic molecular scaffolds in forms that enable life.

Those nitrogen atoms are in molecules that build DNA and RNA. They are present in the backbone peptide bonds of all proteins. They make up the polar head groups of phospholipids that form cell membranes. They are crucial to the functions of seven of the eight B vitamins. They are built in to cellular second messengers as diverse as cyclic AMP and cyclic GMP vs. the inorganic nitric oxide radical.

The secret to unlocking the inert nitrogen atoms lies far back in evolution when anaerobic bacteria, anaerobic because the atmosphere in the beginning was essentially nitrogen gas, evolved an enzyme, nitrogenase, that could reduce dissolved N_2 gas by six electrons to ammonia, NH_3 . Ammonia is a form of nitrogen that living organisms can work with. It is both basic in reaction with protons and nucleophilic in reaction with carbonyl groups of organic metabolites. Indeed, bacterially generated NH_3 is the crossover molecule between inorganic nitrogen chemical biology and organic nitrogen chemical biology.

We use the term nitrogen chemical biology to evaluate how the chemistry of ammonia and its myriad derivatives and downstream metabolites both constrain and enable all the biological phenomena listed above. There are relatively few, mechanistically straightforward nitrogen-containing functional groups that organisms can build. They differ whether one is looking at the inorganic nitrogen chemical biology practiced by the descendants of the single-celled bacteria that learned how to free up the nitrogen chemistry locked in N_2 or at the organic nitrogen chemistry practiced by multicellular eukaryotes, such as humans.

Bacterial inorganic nitrogen metabolism undergirds the biogeochemical cycling of N_2 from atmosphere to organisms and back to atmosphere. The anchor molecules are N_2 , NH_3 , NO_2^- and NO_3^- . The metabolic connections are all reduction–oxidation (redox) processes, electron transfers that support energy metabolism, allowing bacteria to grow on ammonia as a sole energy source or nitrate ions as an energy sink. These bacterial manipulations account for a flux of tens of thousands of tons of N_2 cycling per year.

In contrast, once NH_3 has crossed over to the organic metabolite universe, the great bulk of nitrogen metabolic chemistry occurs without redox changes at nitrogen. Instead, organisms make and interconvert amines and amides, carbamyl groups and ureas, amidines and guanidines. Amino acids are the building blocks for all the proteins that execute the instructions in messenger RNAs. These common nitrogen

functional groups can be embedded in linear organic frameworks or assembled into a diverse array of nitrogen heterocycles, fashioned by simple chemical steps in cells.

The apotheosis of organic nitrogen chemical biology may be in the construction and biological function of those heterocycles. Most notable are the pyrimidines and purines, the building blocks for all DNAs and RNAs across evolution. Those nucleobases offer amines, amides, amidines and guanidines as part structures that make up the hydrogen bonds that lead to A-T and G-C pairs that stabilize DNA double helices and allow the storage of all genomic information in DNA molecules.

Crucial nitrogen heterocycles occur in protein side chains: imidazole rings of histidines and indole rings of tryptophan residues. Seven of the eight essential B vitamins, B₁-B₁₂, provide distinct nitrogen heterocycles to function as coenzymes to enable enzymes to carry out the main reactions of life. A third set of nitrogen heterocycles are the families of tetrapyrrole macrocycles, often termed the pigments of life. They include the red-brown iron-containing hemes, green magnesium chlorophylls, salmon pink colored cobalt-B₁₂ and the yellow nickel-based tetrapyrrole coenzyme in methanogenic bacteria.

This volume categorizes the many life-sustaining roles of nitrogen atoms across the diaspora of metabolites, both low molecular weight and macromolecular biopolymers, based on the chemistry available to nitrogen under physiological conditions. This approach starts with the trivalency of nitrogen with its seven electrons, and the accessible oxidation states from $N = -3$ in NH_3 to $N = +5$ in NO_3^- . It centers on the nucleophilicity and basicity of NH_3 and congeneric organic primary amines, $R-NH_2$, and the simple reactions to form imines and amides. From that point a universe of organic nitrogen chemical biology opens up that follows simple chemical rules to populate the molecules of life.

This book was written to examine nitrogen chemical biology from this molecular perspective. The biochemistry available to nitrogen, as a nucleophile and base catalyst, in organismal metabolism provides an underlying unity to understand and appreciate the special roles that molecules with embedded nitrogens play in the interconnected pathways and structures of life.

Christopher T. Walsh

Dedicated To:

Diana
Allison
Thomas
Sean

Contents

Section I

Chapter 1 Introduction to Nitrogen Chemical Biology

- 1.1 Nitrogen Chemical Biology
- 1.2 Properties of Atomic and Molecular Nitrogen Relevant to Biology
- 1.3 Accessible Nitrogen Oxidation States in Biology
- 1.4 Nitrogen Biogeochemical Cycling 1
- 1.5 Inorganic Nitrogen Metabolism uses the Eight-electron Nitrogen Redox Traverse in Either Direction
- 1.6 Nitrogen Biogeochemical Cycling 2
- 1.7 Crossover to Organic Nitrogen Chemical Biology by Enzymatic Capture of Ammonia
- 1.8 What Kinds of Nitrogen Molecules and Functional Groups Operate in Metabolism?
- 1.9 Nitrogen Heterocycles as Central Structural and Catalytic Groups
- 1.10 One-electron vs. Two-electron Nitrogen Redox Enzymology
- 1.11 Nitrogen Proteomics: Catalysis vs. Posttranslational Modifications
- 1.12 Overview
- References

Section II

Chapter 2 Inorganic Nitrogen Biochemistry and Dinitrogen Reduction to Ammonia by Nitrogenase

- 2.1 Introduction to N_2 , NH_3 and NO_3^-
- 2.2 Nitrification, Denitrification and Reductive Assimilation
- 2.3 N_2 to NH_3 : Biological Fixation of Dinitrogen by Bacterial Nitrogenases
 - 2.3.1 Alternative to Nitrogenase
 - 2.3.2 The Haber–Bosch Industrial Process
- 2.4 Nitrogen-fixing Bacteria
 - 2.4.1 Nitrogenase Structures and Reaction Stoichiometries
 - 2.4.2 Nitrogenase Redox Cofactor Inventories
 - 2.4.3 Thoughts on Nitrogenase Mechanisms
- 2.5 Nitrogen Hydride Biochemistry
- References

Chapter 3 Ammonia Oxidation: Nitrification vs. Anammox

- 3.1 Nitrification: Ammonia Oxidation to Nitrate
 - 3.1.1 Ammonia Oxidase
 - 3.1.2 Hydroxylamine Oxidoreductase
 - 3.1.3 Nitric Oxide Oxidase
 - 3.1.4 Nitrite Oxidase (in Nitrite Oxidizing Bacteria)
 - 3.1.5 Overview: Nitrification: Ammonia to Nitrate Oxidative Pathway
- 3.2 Anammox by the Anaerobic Bacterial Phylum Planctomycetes
 - 3.2.1 Hydrazine Synthase
 - 3.2.2 Hydrazine Dehydrogenase
- 3.3 Ammonification: Degradative Conversion of Organic Nitrogen Metabolites to NH_3
- References

Chapter 4 Nitrate Ion Reductive Metabolism: Bifurcation Between Five-electron or Eight-electron Reductions

- 4.1 Two Modes of Biological Reduction of Nitrate Ions
- 4.2 Nitrate Reduction: NO_3^- to NH_3 for Assimilation: Eight-electron Reduction
 - 4.2.1 Nitrate Reductases
 - 4.2.2 Nitrite Reductases
- 4.3 Denitrification: two NO_3^- to N_2 , Net Ten-electron Reduction
 - 4.3.1 Denitrifying Nitrate Reductase
 - 4.3.2 Denitrifying Nitrite Reductases: One-electron Reductions
 - 4.3.3 Nitric Oxide Reductase in Denitrifying Bacteria
 - 4.3.4 Nitrous Oxide Reductase
- 4.4 Overview of the Metallocenter Inventory in the Denitrification Pathway
- References

Chapter 5 Chemical Strategies Underlying the Biological Flux of Inorganic Nitrogen Species

- 5.1 Traversing the Full Range of Nitrogen Redox States
- 5.2 Three Reductive Metabolic Pathways, One Oxidative Pathway and One Mixed Pathway for Nitrogen Inorganics
- 5.3 Nitrogen Oxide vs. Nitrogen Hydride Chemical Biology
- 5.4 From One to Six Electrons Transferred in a Given Enzyme Site
- 5.5 Redox Cofactor Inventories
- 5.6 Turning to Organic Nitrogen Chemical Biology
- References

Section III

Chapter 6 Amination Reactions: Incorporation of Ammonia Into Amino Acid Frameworks

- 6.1 Enzymatic Amination to Capture Ammonia: Imines and Amides
 - 6.1.1 Glutamate Dehydrogenase: Reductive Amination of Substrate Carbonyl Groups
 - 6.1.2 Glutamine Synthetase: NH_3 Organification via Amide Linkage
 - 6.1.3 Glutamine Oxoglutarate Aminotransferase
- 6.2 Glutamine as Carrier of Latent Ammonia as a General Strategy in Enzymatic Aminations

- 6.2.1 Asparagine and Aspartate vs. Glutamate and Glutamine
- 6.3 Moving Amino Groups Around on Organic Scaffolds
 - 6.3.1 Transaminase Logic
- 6.4 Reactions that Release Ammonia
 - 6.4.1 Amino Acid Deaminases and Ammonia Lyases
 - 6.4.2 Deamination by Amino Acid Dehydrogenases and Oxidases
 - 6.4.3 Amidino Group Deaminases: Purines, Pyrimidines and Nucleic Acids
 - 6.4.4 Release of NH₃ During Tetrapyrrole Macrocyclic Biosynthesis
- 6.5 Overview
- References

Chapter 7 When Ammonia Becomes Toxic: Ureido and Guanidinium Forms of Organic Nitrogen

- 7.1 Maintaining a Steady State Nontoxic Level of Ammonia
- 7.2 Amidines, Guanidines and Ureas
- 7.3 Carbamate, Carbamyl Phosphate and Carbamyl Transfers
 - 7.3.1 Carbamic Acid and Carbamyl Phosphate
 - 7.3.2 Aspartate Transcarbamylase and Ornithine Transcarbamylase
 - 7.3.3 Citrulline to Arginine: Ureido to Guanidinium Group via Addition–Elimination of Aspartate C2-Amino Group
 - 7.3.4 Arginase Completes the Urea Cycle
 - 7.3.5 Amidino Transfer from Arginine to Creatine
- 7.4 Overview
- References

Chapter 8 Metabolic Capture of Amines, Imines and Enamines

- 8.1 Amines, Imines and Enamines
 - 8.1.1 Capture of Substrate Amino Groups in Electrophilic Group Transfer Reactions
 - 8.1.2 Amine Acylations
 - 8.1.3 Metabolic Amine Alkylations
 - 8.1.4 Amine Glycosylations
 - 8.1.5 N-Sulfurylations and N-Phosphorylations
 - 8.1.6 N-Hydroxylations
- 8.2 Imine–Iminium Ion Reactions to Form or Cleave C–C Bonds
 - 8.2.1 Ornithine to Cocaine: Four-iminium Intermediates
 - 8.2.2 Tyrosine to Tetrahydroisoquinoline Alkaloid Framework
- 8.3 Enamine Intermediates as Carbanion Equivalents for C–C Bond Formation
- References

Chapter 9 Nitrogen Heterocycles

- 9.1 Introduction to Nitrogen Heterocycles
- 9.2 Amides, Amidines and Guanidines as H-bond Participants in DNA Base Pairs
- 9.3 Parsing the Origin of Nitrogens in Pyrimidines and Purines Reveals Simple Assembly Logic
 - 9.3.1 Two Nitrogens in Uridine, Three in Cytidine
 - 9.3.2 Five Nitrogens in AMP and GMP
- 9.4 Cyclic Amines, Amides and Other Nitrogen Heterocycles
- 9.5 Tryptophan and Tryptamine as Latent C3 Carbanion–Enamine Equivalent

- 9.5.1 Tryptamine: Acylation, Alkylation and Iminium Formation
- 9.5.2 Physostigmine Biosynthetic Strategy
- 9.5.3 Imine to Beta Carboline Scaffold
- 9.5.4 Tryptophan: Diketopiperazine Formations and Internal Cyclizations
- 9.5.5 Alternate Regiochemistry of Indole Prenylations
- 9.6 Overview
- References

Chapter 10 B Vitamin Nitrogen Heterocycles

- 10.1 Seven Human B Vitamins Contain Functional Nitrogen Heterocycles
- 10.2 The Quaternized Thiazolium Ring in Thiamin-PP the Coenzyme Form of Vitamin B1
- 10.3 The Tricyclic Ring System of Flavin (B2) Coenzymes: Nature's Step-down Redox Transformer
- 10.4 Dihydronicotinamides in Coenzyme Forms of Vitamin B3: Obligate Hydride Transfer Agents
- 10.5 Pyridoxal Phosphate (B6 Coenzyme): a Pyridine for all Amino Acid Transformations
- 10.6 Biotin and Tetrahydrofolate: Carriers of One-carbon Units at all Four Oxidation States
 - 10.6.1 Biotin
 - 10.6.2 Folate
- 10.7 Vitamin B12: Two Organometallic Coenzyme Forms in a Modified Tetrapyrrole Macrocycle
- 10.8 Overview
- References

Chapter 11 Tetrapyrrole Macrocycles: Aminopyrroles and the Pigments of Life

- 11.1 Four Classes of Metallated Tetrapyrrole Macrocycles
- 11.2 Heme Biosynthesis
 - 11.2.1 Two Routes to Aminolevulinate
 - 11.2.2 Aminolevulinate to Porphobilinogen: Imine and Enamine Chemistry
 - 11.2.3 Porphobilinogen
 - 11.2.4 Porphobilinogen To Hydroxybilane
 - 11.2.5 Uroporphyrinogen-III Synthase
 - 11.2.6 Uroporphyrinogen-III to Protoporphyrin IX
- 11.3 Competing Chelatases: Insertion of Fe^{2+} or Mg^{2+}
- 11.4 Two Biosynthetic Branchpoints Yield Tetrapyrrole Macrocycles with Equatorial Iron, Magnesium, Cobalt or Nickel
 - 11.4.1 Branchpoint from Uroporphyrinogen-III to Siroheme, B12, and F430
 - 11.4.2 Branchpoint at Protoporphyrin IX to Chlorophylls
- 11.5 Overview
- References

Chapter 12 Organic Nitrogen Oxygenations

- 12.1 Iron-based vs. Flavin-based N-Oxygenases
- 12.2 Two-electron Oxygenation of Amines to N-Hydroxylamines
 - 12.2.1 An Unusual N–O–Sugar Linkage in the Eneidyne Antitumor

Antibiotic Calicheamicin

- 12.2.2 Lysine to N-Hydroxy-pipecolate in Plant Defense
- 12.3 Six-electron Oxidations of Amines to Nitroso and Nitro Groups
- 12.4 Nitrones and Oximes
- 12.5 Glucosinolates
- 12.6 Ornithine and Lysine to Hydrazine Derivatives via N–O Linkages
- 12.7 Arginine N-Hydroxylation: to D-Cycloserine
- 12.8 Nitric Oxide Synthase. Three-electron N-Oxidation of Arginine Guanidium Group
 - 12.8.1 Four Coenzymes Act As Seriatim to Funnel Electrons to a P450 Type Heme
- 12.9 Human Flavin Monooxygenase (FMO) Isozymes
- References

Chapter 13 Nitriles, Cyanohydrins and Cyanogenesis

- 13.1 Hydrogen Cyanide and Cyanide Anion
- 13.2 Plant Cytochrome P450s Convert Amino Acids to Cyanohydrin Glycosides
 - 13.2.1 The Subclass of Cyp79 Heme N-Oxygenases Convert Amino Acids to Aldoximes
 - 13.2.2 From Aldoxime to Trapped Cyanohydrin Acetals
- 13.3 Two-step Cyanogenesis from the Cyanogenic Glycosides
- 13.4 Recapture of Cyanide by Cyanoalanine Synthase and Conversion to Asparagine or Aspartate
- 13.5 Toyocamycin: Nitrile Formation and Reduction
- 13.6 Additional Cyanogenesis Strategies
 - 13.6.1 Enzymatic Oxidative Fragmentation of 1-Amino-1-carboxy-cyclopropane to Ethylene, Cyanide Ion and Carbon Dioxide
 - 13.6.2 Two Bacterial Amino Acid Fragmentations to Cyanide
- 13.7 Biological Formation of Isonitriles
- 13.8 Overview
- References

Chapter 14 One-electron Nitrogen Chemical Biology

- 14.1 One-electron Transfers To and From Inorganic Nitrogen Atoms
- 14.2 Two-electron vs. One-electron Redox Manifolds on Organic Nitrogen Scaffolds
 - 14.2.1 PLP-dependent Transaminases are Redox Catalysts but Leave the Nitrogen Oxidation State Unchanged
- 14.3 Organic Nitrogen Radical Metabolites
 - 14.3.1 Parallel One- and Two-electron Strategies in Substrate Amine N-Oxygenations
 - 14.3.2 One-electron Transfer in All Phases of Cytochrome P450 N-oxygenations
 - 14.3.3 Nitrogen Radical Coupling to Carbon and Nitrene Transfer by Cytochrome P450s
 - 14.3.4 Cyanogenesis Involves Nitrogen Radicals in Cyanide Ion Formation
 - 14.3.5 Nitric Oxide Synthase Revisited for One-electron Transfer Mechanisms
- 14.4 Nitric Oxide Biology
 - 14.4.1 Direct and Indirect Effects of Nitric Oxide in Higher Eukaryotes

Section IV

Chapter 15 Peptide Bond Formation Strategies in Nonribosomal and Ribosomal Assembly Lines

- 15.1 Peptide Bonds
- 15.2 Strategies for Amino Acid Monomer Activation for Peptide Bond Formation
 - 15.2.1 Conversion of Ammonium Cations to Amine Nucleophiles by Deprotonation
 - 15.2.2 Carboxylate Activation via Mixed Acyl–Phosphoric Anhydrides at the Expense of ATP Side Chain Cleavage
 - 15.2.3 Harvesting the Thermodynamic Activation of Mixed Acyl–Phosphoric Anhydrides in Kinetically Stable Acyl Derivatives
- 15.3 ATP Grasp Family of ADP-forming Amide Bond Synthetases
- 15.4 Peptide Bond Initiation, Elongation and Termination at the Ribosomal RNA Machine
 - 15.4.1 Ribosomal Initiation of Peptide Chain Formation
 - 15.4.2 Peptidyl Chain Elongation at the Ribosome
 - 15.4.3 Peptidyl Chain Termination at the Ribosome
- 15.5 Chemical Logic for Nonribosomal Peptide Synthetase Assembly Lines
 - 15.5.1 NRPS Assembly Line Architecture
- 15.6 ACV Synthetase: A Three Module NRPS Assembly Line
 - 15.6.1 ACV Synthetase Initiation Module
 - 15.6.2 ACV Synthetase Elongation Module
 - 15.6.3 ACV Synthetase Termination Module
- 15.7 Nonribosomal Peptide Variants
- 15.8 Overview
- References

Chapter 16 Nitrogen Side Chains in Enzyme Catalysis and Posttranslational Modifications

- 16.1 Six Proteinogenic Amino Acid Building Blocks with Side Chain Nitrogen Atoms
- 16.2 Transient vs. Stable Covalent Adducts on Protein Residue Nitrogen Side Chains
- 16.3 Histidine as Acid–Base and Nucleophilic Catalyst
 - 16.3.1 Base–Acid Catalysis in Serine and Cysteine Proteases
 - 16.3.2 Histidine as a Nucleophile: Histidine Autokinases
 - 16.3.3 Posttranslational Methylation at N1 or N3 of the Imidazole Ring
- 16.4 Protein Lysine N6 as a Nucleophile for Alkylations, Acylations, AMPylations and Imine–Enamine Tautomerizations
 - 16.4.1 Iminium Ion Formations
 - 16.4.2 AMPylations in DNA Ligase Catalytic Repair of DNA Strand Breaks
 - 16.4.3 Iterative Methylations
 - 16.4.4 (Reversible) Posttranslational Acylations by Acyl CoA Metabolites
 - 16.4.5 Acylations by Information-rich Ubiquitin and SUMO Isoforms
- 16.5 Arginine, an Unlikely Protein Side Chain Nucleophile

- 16.5.1 Bacterial Arginine Kinases
- 16.5.2 Arginine Guanidinium Methylations
- 16.5.3 Protein Arginine Deiminases
- 16.5.4 Arginine ADP-ribosylations
- 16.5.5 Nitroarginine Formation
- 16.6 Asparagine, an Even More Unlikely Protein Side Chain Nucleophile
- 16.7 Posttranslational Glutamine Deamidations
- 16.8 Tryptophan C2-Mannosylation as an Expression of Indole Ring Reactivity
- 16.9 Overview
- References

Subject Index

Section I

This section is composed of Chapter 1 and serves as an introduction to nitrogen chemical biology, across both inorganic and organic nitrogen metabolites. An adult human may contain up to 5 kilograms of nitrogen atoms. That inventory encompasses all the DNA and RNA molecules. The DNA content in a human is a 3 billion base pair genome in every cell, multiplied by approximately 4×10^{13} cells, for a total of approximately 2.4×10^{23} molecules of the purine and pyrimidine building blocks. The nitrogen content also includes all the proteins (up to approximately 25 000 in a fully expressed proteome) in each of the approximately 40 trillion human cells in the body. Then, there are the seven B vitamins that have nitrogen heterocycles crucial for sustaining life, the nitrogen containing phospholipids in their hundreds of millions in every cell and the panoply of low molecular weight amino acids, neurotransmitter amines and other nitrogenous metabolites, including nitric oxide radicals as signaling entities.

On the one hand, nitrogen is the most abundant chemical element on and near the surface of the earth, since N_2 gas molecules comprise 78% of the atmosphere, approaching 4000 trillion tons *in toto*. On the other hand, dinitrogen is an inert, tasteless, gas, inhaled and exhaled with every breath to no good or bad effect (at least at atmospheric pressure). Despite this staggering abundance, none of that dinitrogen is in a form available for use in organismal biology.

One of the solved conundrums of life was the evolution by anaerobic microbes of the enzyme nitrogenase that reduces N_2 by six electrons to NH_3 , an inorganic metabolite that organisms can work with, opening up a rich inorganic and organic nitrogen biology. A crucial feature of nitrogen chemistry, put to full use in its biology, is its ability to undergo an eight-electron redox traverse. The oxidation state of N_2 has both nitrogens at the $N = 0$ oxidation state. Ammonia has N at $N = -3$. Ammonia is oxidized by certain aerobic bacteria all the way to nitrate ions ($N = +5$). We examine inorganic nitrogen chemical biology and the biogeochemical cycling of N_2 to NH_3 to NO_3^- and back to N_2 . Analysis of the underlying chemical logic of those lifegiving redox transformations and the cofactor machinery indicate the key roles that had to be evolved to support those electron transfers in both oxidative and reductive roles.

Ammonia is the most important inorganic nitrogen metabolite. It crosses over to organic metabolism by two enzymatic reactions that fix ammonia into organic scaffolds: (1) imine formation and subsequent reductive amination to glutamate or (2) amide formation as glutamate is converted to glutamine. Most of the subsequent nitrogen chemical biology, to make the additional proteinogenic amino acids and to make the DNA and RNA building blocks, proceeds from amino acids by amination, amide formation, amidine formation, carbamylation of amines to ureas and then amine addition to form guanidine groups. Those central nitrogen groupings occur in both linear metabolites and as the hydrogen bonding structures in purine and pyrimidine heterocyclic base pairs in DNA and RNA. The covalent linkage in proteins joining the condensed amino acid residues head to tail are the amide or peptide bonds.

Nitrogen heterocycles evolved early in life and the molecular logic for the construction of purines and pyrimidines is appropriately simple from amine chemical biology. Not only the purine and pyrimidine heterocycles, the essential building blocks for nucleic acids, but seven of the eight B vitamins each have distinctive nitrogen heterocycles that, separately and in aggregate, enable much of the primary biosynthetic metabolic transformations of life. Particularly noticeable are the so-called pigments of life, the tetrapyrrole macrocycles that include red-brown heme variants, purple siroheme, green chlorophylls, pink vitamin B₁₂ and yellow F430 in methanogenic bacteria.

All these macrocycles are assembled from the aminomethyl pyrrole monomer porphobilinogen, in turn assembled from the amino acid aminolevulinate. The chemistry of the aminomethyl pyrrole monomer as both electrophile and nucleophile in controlled polymerizations is an elegant chapter in nitrogen heterocyclic chemical biology.

Chapter 1 thus outlines the chemical properties of nitrogen that enable it to serve its multiple essential functions in life. The inorganic nitrogen chemical biology, at the heart of nitrogen biogeochemical cycles, is essential to life on the planet by way of initial reduction of N₂ to NH₃. The chemistry then available to ammonia (N = -3) as both base and nucleophile enabled it to cross over into organic chemical biology where it has powered nucleic acid and protein biogenesis strategies and created an interconnected series of nitrogen functional groups that enable both linear and heterocyclic metabolite constructions that create nature's central molecular and information inventory.

Introduction to Nitrogen Chemical Biology

1.1 Nitrogen Chemical Biology

Nitrogen is one of the seven prominent chemical elements in the human body. In decreasing order of abundance, they are oxygen > carbon > hydrogen > nitrogen > calcium > phosphorus > sulfur. The oxygen and hydrogen contents are high because about 50% of the mass of a 70 kg human adult is water. Total body nitrogen is about 5 kg or 11 pounds (In plants nitrogen in all its combined forms constitutes approximately 4% of dry weight).

Within the body's 5 kg are all the nitrogen heterocyclic bases of DNA, approximately 6 billion purine plus pyrimidine bases in each cell multiplied by approximately 4×10^{13} human cells. Also subsumed in that number are all the peptide bonds in whatever fraction of the potential 25 000 proteins of the human proteome are being expressed in each of the 40 trillion or so cells in the body. Also included are a subset of the nitrogenous low molecular weight metabolites that populate so many biosynthetic and degradative metabolic pathways (Figure 1.1).

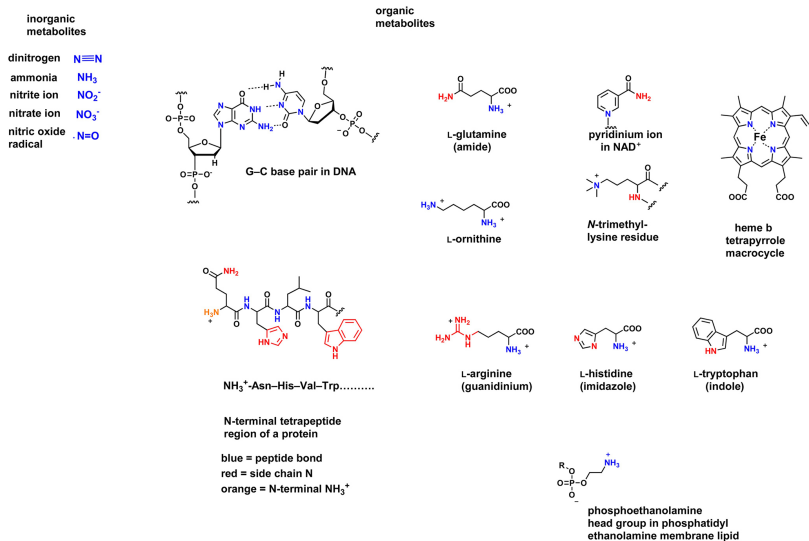


Figure 1.1 Representative inorganic and organic nitrogen metabolites. The five most common inorganic nitrogen metabolites are in the left-hand column. Representative organic nitrogen building blocks for DNA, for proteins and for nicotinamide and tetrapyrrole cofactors are shown in the rest of the figure. Nitrogen atoms are shown in blue.

The left column of [Figure 1.1](#) lists five common inorganic metabolites of nitrogen, containing N–H or N–O bonds but no carbon atoms. The organic nitrogen scaffolds in the right portion of the figure contain different nitrogen functional groups embedded in organic, carbon-based molecular frameworks. We take up inorganic and organic nitrogen metabolites in Sections II and III of this book, respectively.

In terms of global abundance, nitrogen far outdistances the other six elements that make up >98.5% of human body mass, albeit in the inert, invisible, tasteless dinitrogen gas (N_2). At 78% of earth's global atmosphere, N_2 is four times more abundant than O_2 at 21%, and argon a far third at approximately 1% abundance. Humans at rest may breathe in 11 000 liters of air per day. That translates to about 9000 liters of N_2 inhaled and exhaled by each of the 7 billion people on the earth each day, with no chemical reactions between the N_2 breathed in and any of the organs, tissues, or cells of humans or other organisms.

One of the essential features of life is how organisms have evolved to convert the highly abundant inert N_2 gas into usable forms for both inorganic chemical biology of nitrogen (nitrogen compounds with N, H and/or O atoms) and organic chemical biology of nitrogen (amines, amino acids, amides, peptides, purines, pyrimidines, pyrroles, the nitrogen heterocycles of six of the eight B vitamins and more).

This is the definition of *nitrogen chemical biology* that we use in this volume: how the different chemistry available to nitrogen, both in terms of formal redox states and in terms of the plethora of nitrogen-containing functional groups has enabled the manifold essential biological functions that sustain life.

Because nitrogen is so intertwined in life processes at the organismal, tissue, cell and molecular level, we do not try to cover all the biological implications, ramifications and details enabled by nitrogen. These include nucleic acid information storage and transmittal, the myriad functions of proteins with their covalent peptide backbones, the regulation of potentially toxic ammonia levels, hundreds of conversions of free amino acids, the use of the indole ring in tryptophan to engender thousands of natural product scaffolds, the role of side chains in enzymatic catalysis, the full scope of approximately 200 posttranslational modifications, and many more. That would easily fill a dozen books of this size.

Our focus then is on the *biology enabled by the chemistry of nitrogen*: the types of nitrogen functional groups, both inorganic and organic, that are found in biologic systems. How they react and for what purpose (energy, catalysis, information transfer, construction of biosynthetic biopolymer categories, signal transduction, etc.). In sum, what kinds of biology are enabled by the specific nitrogen chemistries.

We focus on the chemistry available to nitrogen in its inorganic and organic metabolite forms under physiological conditions. This gives insight into the reactivity of N_2 , NH_3 and NO_3^- for enzyme-mediated electron transfer in and out of the nitrogen atoms in a lithotrophic energy metabolism that stretches back to early microbial evolution logic. NH_3 is the bridge from inorganic to organic nitrogen chemical biology. On fixation as the amino group in glutamate or the amide group in glutamine, a diaspora of organic nitrogen spreads to a broad swath of metabolic pathways, including protein and DNA and RNA biosynthesis. The simple execution of chemical logic for conversion of organic amines to imines, amides, amidines, ureas and guanidines, in both acyclic and heterocyclic scaffolds, accounts for the essentiality of nitrogen functional groups to life.

In this volume we examine the chemical biology of inorganic nitrogen molecules

first and dissect the logic and mechanisms for microbial interconversion of the “big three”: dinitrogen gas (N_2), gaseous ammonia (NH_3) and dissolved nitrate anions (NO_3^-).

Nitrogen in different molecular forms undergoes cycling on a global scale between atmospheric N_2 , microbial six-electron reduction to NH_3 and subsequent eight-electron oxidation to nitrate by aerobic microorganisms. [1·3](#) In turn, nitrate anions can be reduced biologically by two pathways. Both can be practiced by different anaerobic bacteria. One route is to use nitrate as a terminal acceptor of metabolic electrons, reducing it, instead of O_2 , the normal terminal electron acceptor in aerobes, by eight electrons all the way back to NH_3 . A second alternate anaerobic route is a reduction of two nitrate ions to N_2 ; each nitrate ion is reduced by five electrons, ten electrons overall, to N_2 , net denitrification.

Then we turn to the bridge between inorganic and organic nitrogen chemical biology. That bridging molecule is NH_3 , potentially toxic because of its high nucleophilic reactivity. Next, we examine common organic metabolite scaffolds to examine the strategies organisms have developed for fixing ammonia into those organic molecular frameworks.

1.2 Properties of Atomic and Molecular Nitrogen Relevant to Biology

Nitrogen was first detected and named as “phlogisticated air” by the Scottish physician Daniel Rutherford in 1772, discovered by the absence of function, the inability to support life after oxygen had been removed from air. Lavoisier named nitrogen gas *azote*, meaning “without life” to distinguish it from the life giving oxygen in air. [4](#) The accepted name nitrogen derives from the Greek roots *nitrum* and *genes*, translatable as soda-giving. It is formed by nuclear fusion in stars at high temperature and from lighter elements. The formation of earth out of interstellar particles and dust was thought to involve solid nitrogen that melted and then gasified to dinitrogen gas as the earth warmed and volcanoes emitted interior N_2 . Solid nitrogen melts at $-209\text{ }^\circ\text{C}$ and boils off as N_2 gas at $-195\text{ }^\circ\text{C}$. There are an estimated 4000 trillion tons of nitrogen gas in the atmosphere, amounting to approximately 2.4×10^{21} molecules of N_2 , inert, tasteless, invisible.

Nitrogen is a first-row element in the periodic table; it has seven electrons with an atomic weight of 14 ([Figure 1.2](#)). The ^{14}N isotope accounts for 99.6% of nitrogen in nature. There are several other minor isotopes, of which ^{15}N is a useful nucleus for biological nuclear magnetic resonance (NMR) studies, including assignment of protein structures.

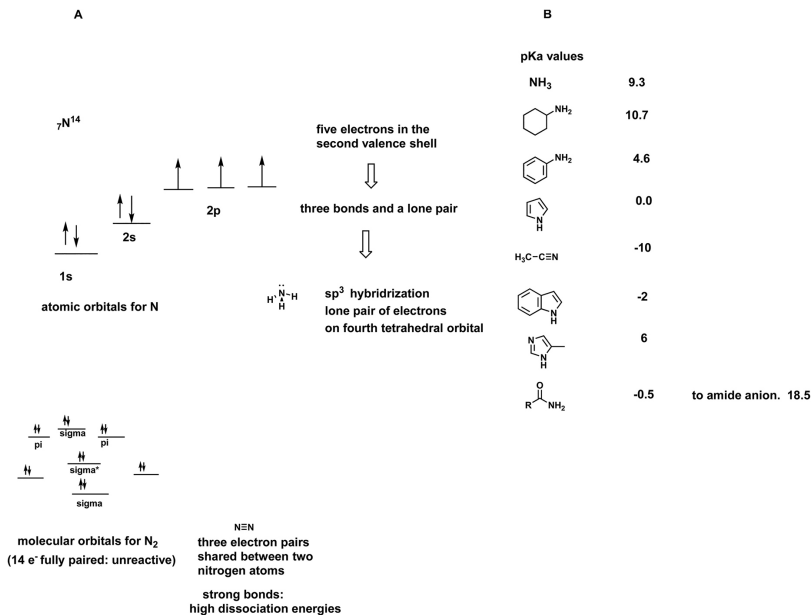


Figure 1.2 (A) Orbital filling diagram for the seven electrons of atomic nitrogen and the molecular orbital diagram for the 14 electrons in N₂. (B) pK_a values for some common nitrogen functional groups (pK_a = pH value where 50% of amine is protonated).

The electron filling scheme for the seven electrons of atomic nitrogen is shown in the atomic orbital diagram to have filled 1s and 2s orbitals. The remaining three electrons fill three 2p orbitals one at a time in the lowest energy configuration. Thus, nitrogen compounds are classically trivalent, such as NH₃. Each hydrogen contributes an electron to each N–H bond. The nitrogen in NH₃ is sp³ hybridized with four tetrahedral orbitals. The fourth orbital is occupied by the nitrogen lone pair.

That lone pair, available in both free ammonia and the organic amine metabolites in cells and organisms, accounts for the reactivity of ammonia and organic amines as both bases and nucleophiles. When reacting as a base, the lone pair on NH₃ reacts with an electrophilic proton, to produce the tetrahedral NH₄⁺ cation. The pH at which there is a 1:1 mix of NH₃ (ammonia) and NH₄⁺ ammonium cation is 9.3. By definition this is the pK_a value. Thus, at physiological pH values near pH 7.3, there will be one part of NH₃ to 99 parts of NH₄⁺.

Nucleophilic reactivity is defined as reaction with an electrophile other than a proton. In biology that could be an electrophilic carbon such as a carbonyl carbon of an aldehyde or ketone, the carbonyl group of an anhydride such as the mixed anhydride between an acyl group and a phosphoric acid, an acyl thioester or an electrophilic phosphorus center (Figure 1.3). The nitrogen lone pair available in amines and ammonia enables them to behave as nucleophiles and give the products shown in Figure 1.3. In particular we will return to imines and amides in one of the later sections of this introductory chapter.

Major Fluxes to Imines and Amides

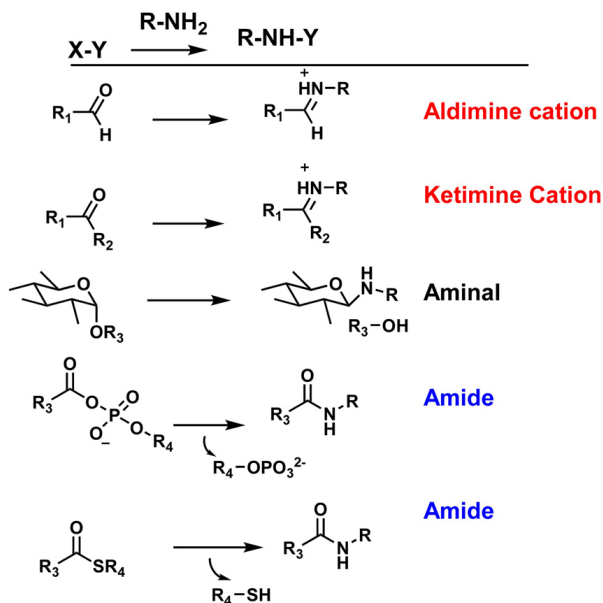
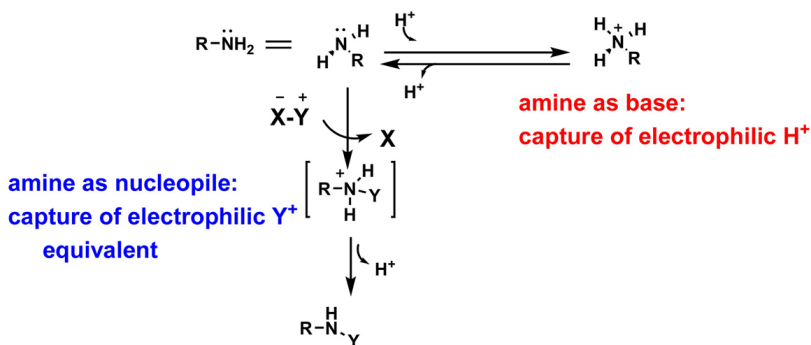


Figure 1.3 Action of amines as bases or nucleophiles: identity of electrophilic group determines outcome. Reaction with proton as electrophiles = action of a base. Reaction with other electrophiles = action as a nucleophile. Typical nucleophilic reactions with cellular electrophiles are with carbonyls at the oxidation state of aldehydes and ketones to generate imines; or with acyl group carbonyls to generate amides.

Figure 1.2 also depicts a molecular orbital diagram for the filling of the $7 + 7 = 14$ electrons in dinitrogen gas. 5 All ground state orbitals are full. There is a strong triple bond, consisting of three shared electron pairs between the two nitrogen atoms. This bond has a very high dissociation energy and accounts for the apparent inertness of

dinitrogen in the atmosphere, towards both abiotic and biotic reagents. One can break the $\text{N}\equiv\text{N}$ triple bond reductively as we shall see in Chapter 2 but it requires both substantial energy input and a mechanism for consecutive six-electron transfers that have a low enough energy barrier to proceed under ambient conditions.

1.3 Accessible Nitrogen Oxidation States in Biology

One of the remarkable facets of nitrogen chemistry played out in its biological roles is the accessibility of nitrogen metabolites that vary over eight formal oxidation states of nitrogen, 6 from -3 in NH_3 , RNH_2 and R-CN (nitrile) metabolites to the $+5$ oxidation state in the major inorganic metabolite nitrate monoanion (NO_3^-) (Figure 1.4).

Inorganic Metabolites: 8 e⁻ Redox Span from N= -3 to N = +5 Oxidation State

N oxidation state	Known metabolites	
-3	NH ₃	ammonia
-2	NH ₂ NH ₂	hydrazine
-1	HN=NH	diazene
0	N ₂	dinitrogen
+1	R-N=O	nitroso
+3	NO ₂ ⁻	nitrite anion
+5	NO ₃ ⁻	nitrate anion

In =contrast.

P stays at +5

Ca stays at +2

S ranges from -2 to +6

Figure 1.4 Accessible redox states of nitrogen in biological metabolites: There is an eight-electron span of oxidation states seen in inorganic nitrogen metabolism, from NH₃ with N at -3, to N₂ with each N = 0, to *nitrite ions* and then *nitrate ions* at N = +3 and = +5, respectively.

By comparison to the other five most common elements in organisms including humans, three of the five are effectively redox inert in terms of stable oxidation states. Oxygen stays in the -2 oxidation state, calcium ions in the $+2$ oxidation state. The overwhelming bulk of hydrogen (e.g. in global water and wherever protons are used) is in the $+1$ oxidation state. [We will see that nicotinamide coenzyme forms of vitamin B3 (Chapter 10) have evolved to act exclusively as hydride transfer agents, with hydrogen at the rare -1 oxidation state.]

Sulfur has an equivalently rich redox traverse as nitrogen. In the case of sulfur this is also an eight-electron traverse, but one electron out of register, from hydrogen sulfide at -2 to sulfate ions (up to 28 mM in global oceans) at the $+6$ oxidation state. In an analogous formal assessment of oxidation states at carbon, the traverse between methane, CH_4 , and carbon dioxide, CO_2 , is an eight-electron oxidation of that carbon atom. The carbon can be viewed as in an oxidation state of -4 in methane (each H at a formal oxidation of $+1$) and as $+4$ in CO_2 (each O at -2). The counting of nitrogen oxidation states is particularly useful for determining when nitrogen is nucleophilic, as in the redox chemistry of inorganic nitrogen compounds in Chapters 2–5.

1.4 Nitrogen Biogeochemical Cycling 1

The concept of biogeochemical cycles is useful for analysis of global flux of elements between different compartments, such as the atmosphere, the oceans and land masses. 3 Cycles for carbon flux, sulfur flux and nitrogen flux are among the ones that biogeologists and climatologists measure and analyze, for example in prediction of global climate changes. There are several ways to visualize the components of the nitrogen biogeochemical cycle. One such diagram is presented in Figure 1.5A. (Note that it leaves out animals, and thereby humans, as qualitatively significant compartments of nitrogen in the biogeochemical cycles.)

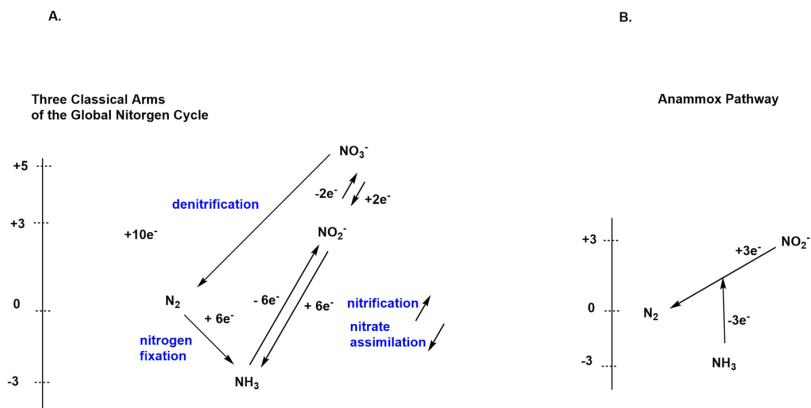


Figure 1.5(A) Cartoon depiction of global cycling of nitrogen, showing the interaction between atmospheric N_2 , soil NH_3 and NO_3^- (nitrate ion) as well as assimilation of nitrate by plants and decomposition of organic nitrogen biomass. Reproduced from ref. 5, <https://en.wikipedia.org/wiki/Nitrogen>, under the terms of the CC BY-SA 3.0 license, <https://creativecommons.org/licenses/by-sa/3.0>. The universe of organic nitrogen metabolites is only hinted at by the word assimilation. (B) The left-hand panel displays the directionality and electron inventory for nitrogen fixation, nitrification or nitrate assimilation and denitrification. The vertical scale indicates the oxidation states of the nitrogen atoms in each

compound. The right-hand panel shows a similar diagram of the anammox arm of nitrogen cycling.

Atmospheric N_2 is reductively metabolized by nitrogen-fixing bacteria, many of which reside symbiotically in legume root nodules, providing NH_3 to those plant roots, while other nitrogen-fixers are free living soil bacteria. ⁷ Given the pK_a of 9.3 for interconversion of NH_3 to NH_4^+ (the pK_a = the pH value where $[NH_3] = [NH_4^+]$), much of the soil ammonia is present as ammonium salts. Nitrifying bacteria were historically thought to comprise two separate clades, those that oxidized ammonia by six electrons to nitrite ions and those that took out another two electrons to produce nitrate ions.

Soil nitrates are a branch point. They can undergo metabolic reduction by denitrifying bacteria, a net conversion to N_2 gas, completing the cycle from N_2 to NH_3 to NO_2^- to NO_3^- and back to N_2 . Alternatively, nitrates can be taken up and assimilated by plants. Not shown under the word *assimilation* in Figure 1.5A is that this process involves reduction of nitrate ions all the way back to ammonia.

These two routes to ammonia, direct reduction of N_2 by nitrogenases or the assimilatory reduction by plants of nitrates to ammonia, are the loci where inorganic nitrogen metabolism crosses over to organic nitrogen metabolism. ³ The diagram in Figure 1.5A shows that trees are more consequential globally than animals and humans in providing organic nitrogenous biomass. The aggregate of all the organic nitrogen compounds in biomass represent a temporary diversion of N_2 that is pulled out of the global cycling. Temporary, because as organisms die, the nitrogenous biomass is degraded by bacteria and fungi all the way down to inorganic NH_3 that brings organic nitrogen metabolites back into the inorganic biogeochemical cycle.

Useful as such biogeochemical cycling diagrams may be in giving a global picture of nitrogen transit, they downplay three key facets of attention in this volume. One is the dramatic change in nitrogen redox states between dinitrogen, ammonia, nitrite and nitrate anions. Those redox state differences are the chemical bases for why nitrogen-fixing bacteria, nitrifying bacteria, denitrifying bacteria, and nitrate assimilatory reducers (including plants) can harvest energy from the chemical transformations between these four inorganic nitrogen molecules. They are discussed in the next section and illustrated in Figure 1.5B.

A second omission from Figure 1.5A is an additional microbial pathway, the anammox pathway, taken up in Chapter 4 and sketched briefly in Figure 1.5B, that contributes significant global oceanic N_2 return by marine microbes. ⁸

Third, and central to the course of this volume, is that biogeochemical diagrams omit the mention of any organic nitrogen compounds. These omissions reflect the geochemical focus on global cycling of elements.

From a nitrogen chemical biology perspective, they ignore all of molecular biology, including all eukaryotic biology. Harking back to the few organic nitrogen molecules in Figure 1.1, those omissions include DNA, RNA, proteins, amino acids, tetrapyrrole macrocycles and the framework of 6–10 nitrogen functional groups that enable the chemistry of life. Those omissions are the subjects of this book as it codifies the underlying logic of nitrogen chemical biology.

1.5 Inorganic Nitrogen Metabolism uses the Eight-electron Nitrogen Redox Traverse in Either Direction

As it happens, microbes learned to use the most positive inorganic forms of sulfur

(sulfate dianions) and nitrogen (nitrate anions) as terminal electron acceptors for metabolically derived electrons to eke out an energy strategy to make a living anaerobically as early denizens on an anaerobic planet. In this sense, the parallel but independent redox versatility of nitrogen and sulfur molecules to engage in facile redox chemical biology was a crucial enablement to energy metabolism in early microbial life on the planet.

Later in evolution as the planet gradually became oxygenated from the actions of photosynthetic cyanobacteria that oxidized water to O_2 on a global scale, microbes that learned to reduce O_2 as terminal electron acceptor could live as chemolithoautotrophs (organisms that grow on inorganic chemicals). Oxidation of ammonia to nitrate ions or in the sulfur oxidizers, hydrogen sulfide to sulfate anions, are both eight-electron traverses. We will see that, to perform the eight-electron redox traverses for energy metabolism in either the reductive or the oxidative direction, the anaerobic and aerobic microbes needed to use redox active metal cofactors such as iron, copper and molybdenum ions in stable coordination scaffolds.

Of the seven metabolites in [Figure 1.3](#) spanning the full range of accessible nitrogen oxidation states, six of the seven are inorganic nitrogen compounds. Only the C-nitroso metabolite has nitrogen in an organic framework. We will see in Chapter 2 that the conversion of N_2 at formal oxidation state zero, goes through diazene and hydrazine equivalents in nitrogenase catalytic cycles to produce two molecules of ammonia. Each nitrogen atom in N_2 has undergone a three-electron reduction. Two nitrogen atoms times three electrons equals a six-electron reduction cycle for nitrogenase.

Analogously, the evolutionarily late-arriving ammonia oxidizing bacteria, taking ammonia all the way to nitrate do so by action of four consecutive enzymes, resulting in the full eight-electron oxidation of nitrogen from its most reduced $N = -3$ state to its most oxidized $N = +5$ oxidation state.

1.6 Nitrogen Biogeochemical Cycling 2

Two other inorganic nitrogen redox pathways contribute to the global flux of nitrogen between the atmosphere, the oceans, the soil, and various forms of biomass ([Figure 1.5B](#)). This is nitrogen chemistry and transport on a truly large scale, thousands of tons per year.

Typically, four to five metabolic pathways are depicted in global cycling of nitrogen atoms, four of them known throughout the 20th century and the fifth, the anammox reaction practiced by anaerobic bacteria, only detected in the past 30 years. [3-8](#) In a real sense planetary life is enabled by the fastidiously anaerobic bacteria, some free living, many symbiotic on legume plants, that carry out the six-electron reduction of abundant but unusable nitrogen in the form of N_2 to two molecules of NH_3 . We note below that NH_3 is the crossover metabolite from inorganic to organic nitrogen chemical biology.

The second pathway is the aerobic microorganism oxidation of ammonia to nitrate, referred to as global nitrification. These bacteria use NH_3 as their energy source, stripping out the electrons and passing them down membrane electron transport chains to O_2 as a terminal acceptor to harvest energy from the eight-electron oxidation of ammonia to nitrate monoanion. In addition, it turns out many plants have evolved to take up nitrate anions as a preferred source of nitrogen.

Plants then practice the third inorganic nitrogen redox pathway, assimilative nitrate reduction. They take nitrate ions all the way back down through the eight-electron traverse, this time putting in eight electrons, *via* different sets of catalysts to get to NH_3 . It may be that plants have evolved to make their own NH_3 from nitrate to avoid the potential toxicity from acutely high levels of NH_3 that can be an uncontrollably

strong nucleophile. Humans likewise keep free ammonia or ammonium cation levels within low limits as we shall explore in Chapter 7.

A fourth redox pathway is an alternate reductive fate for nitrate ions. Rather than a full eight-electron reductive traverse of nitrate to ammonia, nitrate ions are instead reductively shunted by picking up five electrons to yield half an equivalent of an N_2 molecule. More accurately, two nitrate ions undergo a net ten-electron reduction to N_2 . The three pathways of dinitrogen fixation, nitrification and denitrification (Figure 1.5A) formally complete the three arms of the nitrogen biogeochemical cycle (N_2 to 2NH_3 ; NH_3 to NO_3^- ; 2NO_3^- to N_2).

However, by the end of the 20th century it was clear that the measured fluxes in these pathways could not account for the actual amount of N_2 returned on a global scale. This led to a successful search for and discovery of slow growing obligate anaerobic bacteria, originally isolated from industrial aqueous sludge, that carry out the fifth redox pathway. ⁸ The anammox (anaerobic ammonia oxidation) pathway combines NH_3 and a nitrite ion to yield N_2 (Figure 1.5B). We examine the logic and enzymatic machinery of what is effectively a shunt, bypassing nitrate ion involvement. Formally, NH_3 is at a nitrogen oxidation state of -3 ; the nitrite ion at an oxidation state of $+3$. Their combination leads to N_2 where each nitrogen has a formal oxidation state of zero. The redox stoichiometry works out, albeit some unusual nitrogen chemical biology is practiced.

In Chapters 2–5, comprising Section II of this volume we take up these five pathways of inorganic nitrogen metabolism highlighted as key for nitrogen global recycling by biological routes. This chemical logic and attendant enzymatic machinery are largely employed by microorganisms using a widely inventive set of iron and copper cofactors to achieve the specific electron transfers. The chemical biology in those chapters is all about nitrogen hydride and nitrogen oxide metabolite reactions.

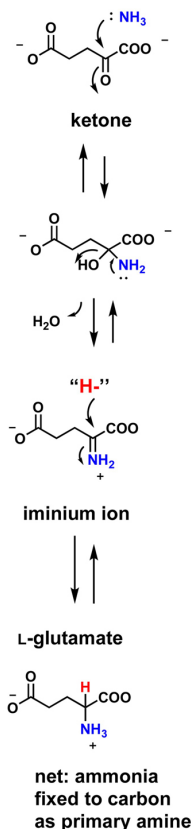
We turn next to a rich and varied nitrogen metabolic universe, the reactions in organic biomass that are usually relegated to a small undistinguished pile in the maps of nitrogen biogeochemical cycling, in part because organic nitrogen compounds are not volatile and do not escape into the atmosphere.

1.7 Crossover to Organic Nitrogen Chemical Biology by Enzymatic Capture of Ammonia

Section III of the book, comprising Chapters 6–14 turns from inorganic nitrogen chemical biology to the metabolic universe of organic nitrogen chemical biology. This is the realm of amino acids and their interconversions. It is the realm of the purine and pyrimidine heterocyclic bases in DNAs and RNAs. It is the realm of the hundreds of thousands of peptide bonds in proteomes. It is the realm of the nitrogen heterocycles in the coenzyme forms of seven of the eight B vitamins.

The crossover metabolite is NH_3 . By far the greatest flux for enzymatic incorporation into organic scaffolds are the reactions creating glutamate and glutamine, respectively (Figure 1.6). The nucleophilic nitrogen in NH_3 undergoes two kinds of reactions with electrophilic carbon centers.

1. Imine Formation then Hydride Reduction:
Reductive Amination



2. Acid + Amine

amide: driven by ATP cleavage

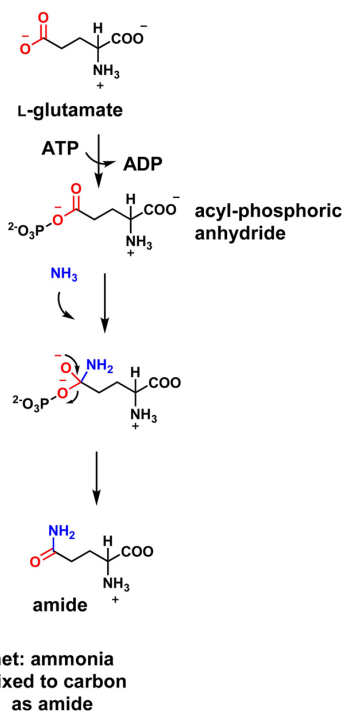


Figure 1.6 Imine and amide formations: NH_3 as nucleophile. NH_3 is the crossover metabolite from inorganic to organic nitrogen chemical biology. NH_3 is captured in imine and then reduced, or captured by ATP-dependent amide formation.

The first is imine or iminium formation with aldehydes and ketones. Iminium adduct formation *via* reversible dehydration is generalized throughout amino acid metabolism. The step that fixes the nitrogen firmly into the original keto acid scaffold is subsequent hydride transfer from the reduced nicotinamide coenzyme (Chapters 6 and 10), turning the hydrolytically reversible iminium ion to the stably bonded amine. This is reductive amination. ⁹

The second mode of fixation of NH_3 into organic frameworks is the enzymatic conversion of the C_5 carboxylate of L-glutamate to the carboxamide, a net amidation reaction⁹ (Figure 1.6). Amides are uphill thermodynamically from the amine and carboxylate fragments and their formation always requires energy input. Almost always, cells spend a molecule of the adenosine triphosphate (ATP) energy currency every time an amide bond is formed. This is the underlying strategy for generating all the covalent peptide bonds in the backbones of all proteins in proteomes.

Examination of glutamine from this crossover perspective reveals that both nitrogens derive from enzymatic fixation of inorganic NH_3 , once by reductive amination *via* the initial iminium adduct (at C_2) and once by amide bond formation at C_5 . These two reactions are consequences of nucleophilic nitrogen seeking out electrophilic carbonyl

groups (Figure 1.6)

1.8 What Kinds of Nitrogen Molecules and Functional Groups Operate in Metabolism?

In contrast to the focus on changes in oxidation state of N–H and N–O compounds in inorganic nitrogen chemical biology, the great bulk of organic nitrogen metabolism involves no change in the formal oxidation states of nitrogen. The most common kinds of transformations are summarized in Figure 1.7A: amines to amides to amidines or amine carbamylation to ureas (ureido groups) and then amination of the ureas to guanidinium groups. One can think of the amidines as imine forms of amides while one can think of guanidines as the imine forms of ureas, tying together these five organic nitrogen functional groups. Clearly, as one proceeds metabolically from amine to amidine, two nitrogen atoms end up in the functional group. The route from amine to urea to guanidine brings together three nitrogen atoms. We explore the molecular logic underlying these biosynthetic transformations in Chapters 7 and 8. One line of evidence that the amide, ureido, amidine and guanidine functional groups offer chemical versatility is shown in Figure 1.7B where the four building blocks of DNA, adenosine, guanosine, cytosine and thymine contain part structures that are amides, amidines, ureas and guanidine groups. The tautomeric equilibrium between amidine isomers and guanidine isomers also sets up different H-bonding possibilities that are the fundamentals rules for base-pairings between A–T and G–C in DNA.

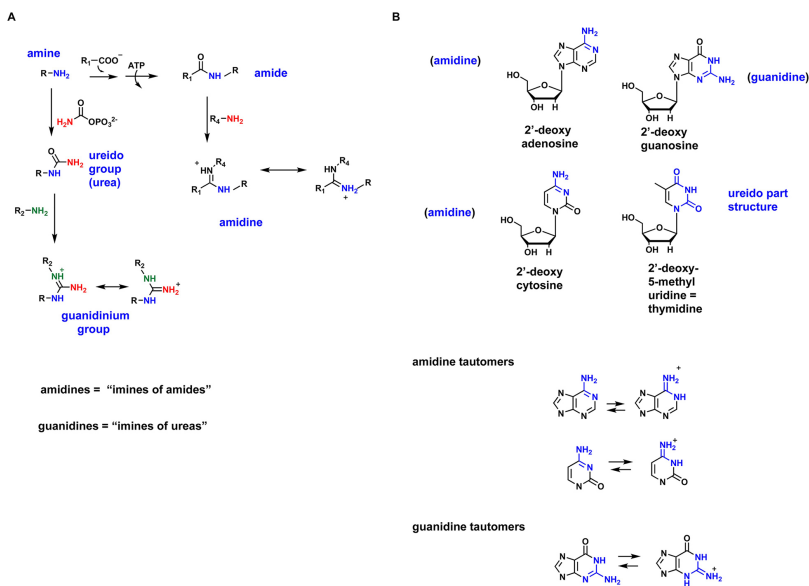


Figure 1.7(A) Five common interconnected functional groups in nitrogen metabolites: amines, amides, amidines, ureas and guanidines; (B) amides, amidines, ureas and guanidines as part structures in the purine and pyrimidine building blocks for DNA and RNA.

1.9 Nitrogen Heterocycles as Central Structural and Catalytic Groups

One can classify the many kinds of nitrogen heterocycles that crop up in distinct areas of organic nitrogen chemical biology into three categories. The first are the above-mentioned heterocycles that store the cellular blueprints in DNA and RNA. These are clearly ancient molecules in evolution. We examine these simple building blocks and molecular logic for installation of the 2- (uracil), 3- (cytosine), 4- (inosine) and 5- (adenine, guanine) nitrogens used throughout all biological species.

A second central category of nitrogen heterocycles are found in the side chains of two of the proteinogenic amino acids and in the construction of tetrapyrrole macrocycles (Figure 1.8A). One amino acid is histidine with its imidazole group. The other amino acid is tryptophan with its bicyclic fused phenylpyrrole ring system, the indole ring. The third member is the pyrrole ring that reveals its chemical virtuosity in the biosynthesis of hemes, chlorophylls, vitamin B₁₂ and the methanogenic coenzyme F₄₃₀, key pigments of life. The combinations of basicity, nucleophilicity, enamine and carbanion reactivity enable reaction types not available to the apoprotein partners.

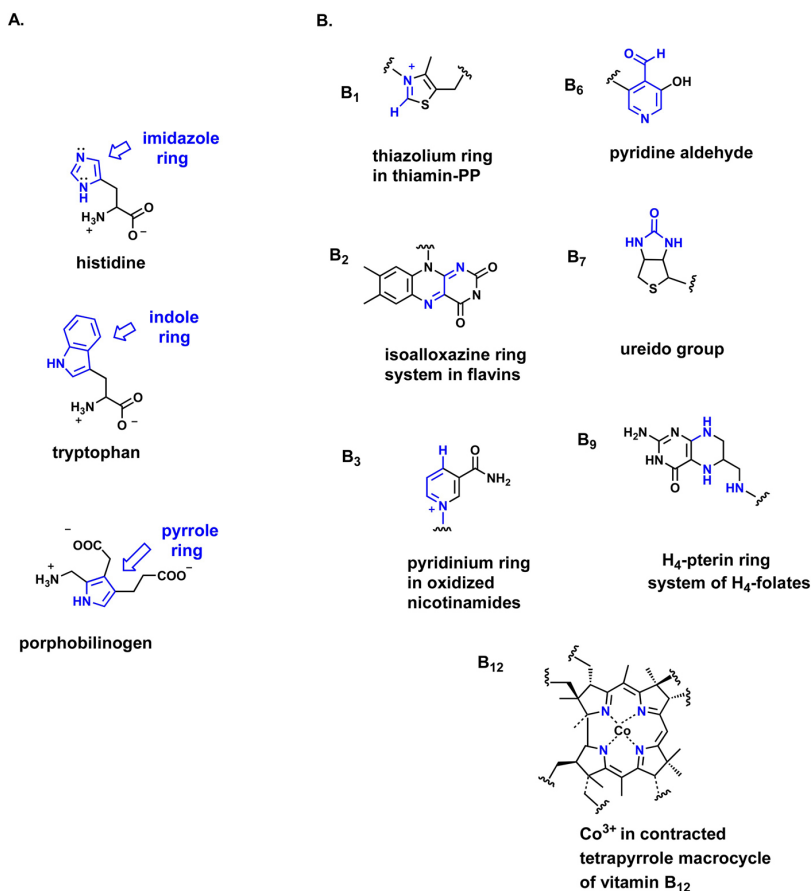


Figure 1.8 (A) Three key nitrogen heterocycle side chains in amino acid metabolism; (B). Seven different nitrogen heterocycles allow the coenzyme forms

The third group of nitrogen heterocycles contains seven of the eight B vitamins (Figure 1.8B). All told, there are 13 human vitamins, nine water-soluble and four (vitamins A, D, E and K) lipid-soluble. Of the water-soluble vitamins eight are grouped as B vitamins and the ninth stands alone as vitamin C. Each of the eight B vitamins function as coenzyme forms to enable classes of metabolically central reactions to proceed rapidly and with specificity.

The active coenzyme forms of seven of the eight B vitamins, B₁ (thiamin), B₂ (riboflavin), B₃ (niacin = nicotinic acid), B₆ (pyridoxine), B₇ (biotin), B₉ (folate) and B₁₂ (cobalamin), contain one or more nitrogen heterocycles that are the functional centers of each coenzyme. 10 The chemical breadth, representing biosynthetic virtuosity by producer microbes early in evolution, is remarkable and tailored to eight different kinds of nitrogen chemical biology transformations, explored in detail in Chapter 10.

1.10 One-electron vs. Two-electron Nitrogen Redox Enzymology

In three of the of inorganic redox pathways of Figure 1.5B, involving inorganic oxides of nitrogen, the odd electron metabolite nitric oxide is an “on pathway” metabolite, Nitric oxide has most of the unpaired electron density on nitrogen (Figure 1.9), qualifying as a nitrogen-centered free radical. Clearly, nitric oxide radicals are stable enough in those microbial cellular environments to serve as intermediate metabolites. Estimates of nitric oxide lifetimes in cells are in the order of 5–10 seconds, long enough to diffuse between enzyme active sites. The enzymes that generate or consume the nitric oxide radical in bacterial metabolism must be able to carry out *one-electron chemistry* on nitrogen atoms. Many of the redox cofactors in the five pathways of Figure 1.5 are heme iron or copper enzymes, obligate one-electron acceptor transition metals.

1. Nitric oxide formed in three inorganic nitrogen redox pathways

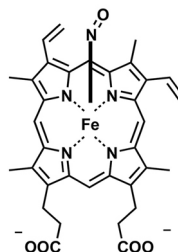
a. denitrification: from nitrite ion



b. anammox: from nitrite ion



c. nitrification: from hydroxylamine



nitric oxide forms tight axial
nitrosyl radical complexes
with Fe^{II} heme

2. Nitric oxide in one organic nitrogen pathway by nitric oxide synthases

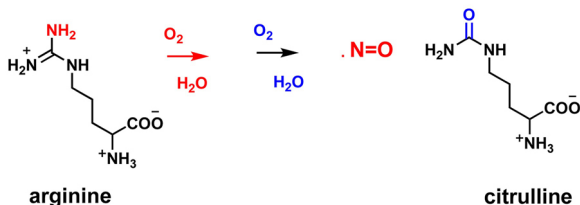
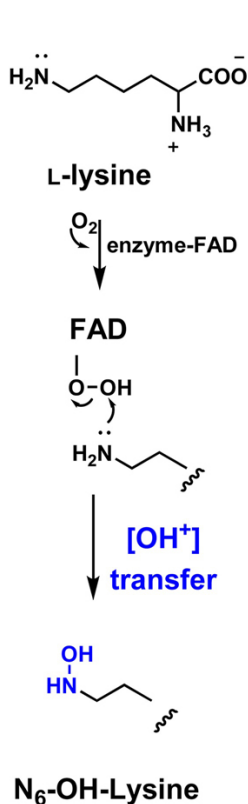


Figure 1.9 The nitrogen oxide radical as a ubiquitous metabolite in prokaryotic inorganic and eukaryotic organic nitrogen chemical biology.

It was a surprise some 30 years ago when humans were observed to use the nitric oxide radical as an endogenous signaling metabolite *via* nitrosyl radical complexation with the regulatory heme iron tetrapyrrole in the guanylate cyclase enzyme (Figure 1.9) that, *inter alia*, regulates blood pressure by smooth muscle relaxation. 11-13 Three isoforms of nitric oxide synthases use arginine and O_2 to produce citrulline along with the nitric oxide radical, a net three-electron oxidation at the guanidinium group of arginine (Figure 1.9), proof that humans also carry out *odd electron oxidations* at nitrogen atoms.

Examination of plant and microbial organic amine oxygenations reveal both one-electron and two-electron reaction manifolds, utilizing different cofactors for reductive activation of cosubstrate molecular oxygen (Figure 1.10). *N*-hydroxylation of amine side chains in ornithine or lysine on the way to production of bacterial siderophores to capture limiting Fe^{III} atoms are two-electron processes using flavin (vitamin B_2 ; Figure 1.8B) coenzymes. In contrast, plants employ heme-based cytochrome P450 *N*-monooxygenases to carry out one-electron chemistry on the C_2 -amine of amino acid substrates. These undergo multiple hydroxylations on the way to oximes, nitroso and nitrile products.

1. FAD-dependent *N*-oxygenases



2. Heme iron-dependent *N*-oxygenases

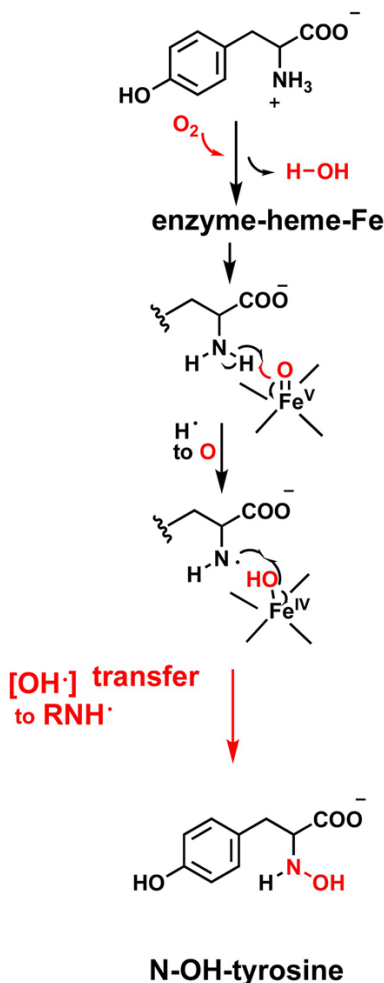


Figure 1.10 Two-electron vs. one-electron reaction manifolds in enzymatic amine *N*-oxygenations: flavoenzymes vs. cytochrome P450s as catalysts that transfer [OH⁺] vs. [OH·] equivalents to amine substrates.

1.11 Nitrogen Proteomics: Catalysis vs. Posttranslational Modifications

Section IV, comprising Chapters 15 and 16, takes a focused view of nitrogen proteomics. The sole covalent linkage in most cytoplasmic proteins, where cysteine (Cys) thiols side chains are reduced, are the peptide bonds that link together the dehydrated amino acid monomers. For an average size protein of 300 residues (33 kDa molecular weight), there are 299 connecting peptide bonds (Figure 1.1) between the (300) building blocks. All amide bond formations are uphill energetically from the

amine and carboxylate components and require ATP cleavage to activate the carboxy component as an electrophilic acyl-phosphoric anhydride. [14](#)

There are two patterns to ATP side chain cleavage: [1](#) to adenosine diphosphate (ADP) and inorganic phosphate (Pi) in soluble nonribosomal peptide bond formations as in glutathione biosynthesis. However, all protein-containing peptide bonds are generated during: [2](#) ATP cleavage to adenosine monophosphate (AMP) and inorganic pyrophosphate (PPi) by way of aminoacyl-AMP intermediates. In turn, there are two kinds of cellular machineries for directing the activated aminoacyl-AMP monomer to peptide binds. The predominant nanomachinery is in the messenger RNA-programmed ribonucleoprotein ribosomes where the peptidyl transferase center is purely an RNA catalyst. The actual molecules polymerized on the ribosome are aminoacyl oxoesters in the form of aminoacyl tRNAs that read triplet mRNA codons. In contrast, nonribosomal protein synthetases in bacteria and fungi are mRNA-independent and use tethered aminoacyl thioesters in place of oxoesters.

Six of the 20 proteinogenic amino acid building blocks have nitrogen atoms in their side chains. Those nitrogen functional groups include the imidazole ring of histidine, the side chain amides in glutamine and asparagine, the side chain N₆-amine of lysine, the guanidinium group of arginine and the indole side chain of tryptophan ([Figure 1.1](#)). Each of these side chain nitrogen functional groups offer different basicity and nucleophilicity options ([Figure 1.3](#)) both for transient participation in enzymatic catalytic cycles and in formation of stable posttranslational covalent derivatives.

The focus on nitrogen side chains in proteinogenic residues illuminates the ways in which the reactivity of those side chains in the free amino acid metabolites are carried over into protein functionalities. Necessarily, they offer an incomplete view of the full range of both enzymatic catalytic groups [Cys, and serine (Ser) residues for two] and of posttranslational modifications^{[15](#)} [Cys, Ser, threonine (Thr) and tyrosine (Tyr), for example], but the two chapters of Section IV do offer insights into the logic and mechanism of nitrogen proteomics.

1.12 Overview

The approach of this volume to nitrogen-centric chemical biology is uniquely integrative. It takes a functional group approach to both the inorganic and organic sides of nitrogen metabolites that are treated as a single unit. We follow the path of ammonia from its formation by N₂ reduction by nitrogenase to its reaction by anammox and nitrification pathways in biogeochemical cycling of nitrogen. We follow ammonia as the sole crossover metabolite for “organification” as it gets incorporated by way of imine and amide linkages into the universe of organic scaffolds. We examine how the commonly accessible nitrogen functional groups enable such disparate biologies as information transfer, nitrogen proteomics, enzymatic catalysis mediated by nitrogen heterocyclic coenzyme forms, phospholipid membranes, DNA repair and signal transductions at both small molecule and protein levels.

In laying out the underlying chemical logic of nitrogen reactivity in biological systems, this approach avoids a descriptive compendium of different biologies. Instead, it seeks to provide a coherent unitary exposition of how the common nitrogen functional groups arise and interconvert from accessible enzymatic transformations. The relatively limited inventory of nitrogen functional groups constrains and enables organismal, cell and molecular biologies. In this context the citations of the original literature are purposely light. Many of the references are to recent articles or reviews that will enable interested readers to delve into the many layers of biology enabled by the underlying chemistry available to inorganic and organic nitrogen metabolites.

References

1. L. A. Bristow, W. Mohr, S. Ahmerkamp and M. M. M. Kuypers Nutrients that limit growth in the ocean, *Curr. Biol.*, 2017, **27**, R432-R510.
2. L. B. Maia and J. Moura How biology handles nitrite, *Chem. Rev.*, 2014, **114**, 5273-5357.
3. D. Fowler, M. Coyle, U. Skiba, M. A. Sutton, J. N. Cape, R. Reis, L. J. Sheppard, A. Jenkins, B. Grizzetti, J. N. Galloway, P. Vitousek, A. Leach, A. F. Bouwman, K. Butterbach-Bahl, F. Dentener, D. Stevenson, V. Amann and M. Voss The global nitrogen cycle in the twenty-first century, *Philos. Trans. R. Soc. London, Ser. B*, 2013, **368**, 20130164.
4. R. Oblack, *Nitrogen: Gases in the Atmosphere*, 2020, www.thoughtco.com.
5. Wikipedia, *Nitrogen*, <https://en.wikipedia.org/wiki/Nitrogen>.
6. N. Connor, *Nitrogen-electronic Configuration and Oxidation States*, 2020, www.periodic-table.org.
7. M. Petruzzello, *Editors of Encyclopedia Britannica, Nitrogen Fixation*, 2020, www.britannica.com.
8. B. Kortal, W. Maalcke, N. M. de Almeida, I. Curpus, J. Gloerich, W. Geerts, H. J. M. O. den Camp, H. R. Harhangi, E. M. Janssen-Megens, K.-J. Francoijs, H. G. Stunnenberg, J. T. Keltjens, M. S. M. Jetten and M. Strous Molecular mechanisms of anaerobic ammonium oxidation, *Nature*, 2011, **479**, 127-130.
9. M. C. Walker and W. van der Donk The many roles of glutamate in metabolism, *J. Ind. Microbiol. Biotechnol.*, 2015, **43**, 419-430.
10. C. T. Walsh and Y. Tang, *The Chemical Biology of Human Vitamins*, Royal Society of Chemistry, London, 2019, p. 446.
11. J. T. Groves and C. C.-Y. Wang Nitric oxide synthase: mechanisms and models, *Curr. Opin. Chem. Biol.*, 2000, **4**, 686-695.
12. D. D. Thomas, L. Ridnour, J. S. Isenberg, W. Flores-Santana, C. H. Switzer, S. Donzellie, P. Hussain, C. Vecoli, N. Paolocci, S. Ambs, C. Colton, C. Harris, D. D. Roberts and D. A. S. Wink The chemical biology of nitric oxide. Implications in cellular signaling, *Free Radical Biol. Med.*, 2008, **45**, 18-31.
13. C. Fang Mechanism of nitric oxide synthase regulation: electron transfer and interdomain interactions, *Coord. Chem. Rev.*, 2012, **256**, 393-411.
14. C. T. Walsh, B. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
15. C. T. Walsh, *Posttranslational Modification of Proteins: Expanding Nature's Inventory*, Roberts & Co, Englewood, Colorado, 2006.

Section II

Section II deals with the chemical biology of inorganic nitrogen compounds. These are simple molecules. Lacking any carbon atoms, these molecules have N-N, or N-H and/or N-O bonds. Four prominent inorganic metabolites are dinitrogen gas (N_2), ammonia (NH_3), nitrite ion (NO_2^-) and nitrate ion (NO_3^-).

Among the notable features of these nitrogen hydrides (ammonia) and nitrogen oxides (nitrite and nitrate anions) is that they span the full oxidation state spectrum of nitrogen. Nitrate is in the +5 oxidation state at nitrogen, and nitrite, two electrons down, at the +3 oxidation state. N_2 , definitionally, has both nitrogens at the $\text{N} = 0$ oxidation state. NH_3 represents nitrogen in its most reduced, most basic, most nucleophilic state with $\text{N} = -3$.

Of the five other chemical elements (carbon, hydrogen, oxygen, sulfur and phosphorus) that comprise >99% of the mass of free-living organisms, only sulfur shows equivalent redox versatility in its chemistry. Indeed, both inorganic sulfur chemical biology and inorganic nitrogen biology utilize these eight-electron redox traverses in parallel strategies, to allow microbes to use the most reduced compounds, NH_3 or H_2S , as sole energy sources. They eke out chemolithotrophic lifestyles as they remove electrons by oxidative enzymatic processes and harvest some fraction of the energy thereby released in molecules that can do useful cellular work.

At the other end of the redox scales, the most oxidized forms of sulfur and nitrogen are, respectively, sulfate ions ($\text{S} = +6$) and nitrate ions ($\text{N} = +5$). Distinct clades of anaerobic bacteria, sulfate reducers vs. nitrate reducers, have evolved enzyme and cofactor electron transfer machinery to use sulfate dianions or nitrate monoanions as the terminal electron acceptors in membrane electron transport. In these anaerobic microbes, nitrate or sulfate replace O_2 normally used by aerobes. Electrons fall through potential drops as sulfate and nitrate are reduced. Some of that energy is used for anisotropic pumping of protons across bacterial membranes to build up a transmembrane electrochemical potential that is ultimately convertible to ATP by membrane-embedded ATP synthetases. Thus, redox reactions at both ends of the nitrogen oxidation states (NH_3 to NO_3^-) and back, are central to inorganic nitrogen chemical biology in energy metabolism.

The metabolic conundrum in the inorganic context is dinitrogen gas, N_2 . Almost infinitely abundant, when considered as the 78% of the global atmosphere, the remarkably strong N-N triple bond in N_2 leaves it chemically unreactive at ambient temperature and pressure. Only in some lightning strikes is enough energy injected naturally to fragment N_2 reductively. Alternatively, the Haber-Bosch process, introduced at an industrial scale around 1910, at very high temperatures and very high pressures will allow H_2 to reduce N_2 over iron surfaces.

When anaerobic cyanobacteria and other bacterial species evolved nitrogenase catalysts to reduce N_2 by six electrons to NH_3 , life could evolve. NH_3 is a workable source of nitrogen. With properties both as a base in reactions with protons and as a nucleophile in reaction with electrophiles other than protons, NH_3 also serves as an

energy source and crossover molecule to open up organic nitrogen chemical biology.

Study of the enzyme nitrogenase in detail has gone on for more than 50 years and some essential secrets remain even though the research literature is immense. The extensive reach of nitrogen chemical biology - both inorganic and organic - permeates all aspects of molecular biology and biochemistry. The biological ramifications enabled by the several relevant topics in nitrogen chemistry are often beyond the scope of this one volume. However, the chemocentric analysis of nitrogen reactivity in its various common metabolites provides a consistent underlying molecular logic that explains and predicts the behavior of amino acids and proteins, nucleic acid building blocks and their condensed polymers RNA and DNA, vitamins and enzymatic cofactors.

Chapters 2-5 explore the four major pathways in biological cycling of global nitrogen, with a focus on the changes in nitrogen redox states and the roles of iron and copper cofactors that enable those enzymatic transformations. Three pathways proceed only in anaerobic microbes, reflecting oxygen sensitivity of the redox-active transition metal cofactors and speak to an evolution of this inorganic nitrogen chemical biology before the global oxygenation of the planet.

Nitrogenase is all about N-H chemical biology as it transmutes ATP hydrolysis energy into six electron reduction of N_2 to two NH_3 molecules. The reaction pathways probably involve bound diazene and hydrazine, but they have not been definitively characterized as competent intermediates, two and four electrons down from N_2 . This enzyme may be the apotheosis of iron-sulfur cluster biochemistry as electron transfer conduits, in both the P (Fe_8S_7) and M clusters (Fe_7S_9CMo).

The second anaerobic pathway is the anammox reaction practiced by marine bacteria, taking NH_3 ($N = -3$) and nitrite ion ($N = +3$) and generating N_2 (both nitrogens at $N = 0$). Ammonia is oxidized as nitrite is reduced, N-N bond formation is established and the pathway switches from N-O to N-H chemistry *via* action of the enzyme hydrazine synthase.

The third anaerobic pathway is denitrification, where a pair of nitrate ions are each reduced by five electrons as N_2 is formed along with six water molecules. This process, along with anammox, are the two pathways that return N_2 to the atmosphere in the biogeochemical nitrogen cycle.

The sole aerobic pathway is practiced by nitrifying bacteria, converting NH_3 to NO_3^- , removing eight electrons as ammonia ($N = -3$) is oxygenated to nitrate ($N = +5$). Nitrifiers can grow on ammonia as sole nitrogen source. Often one sees synergy between microbes that stop at the nitrite ion (net six-electron oxidation) and nitrate producers that take nitrite the last two electrons up to nitrate.

During the course of evolution of these inorganic nitrogen redox pathways microbes have invented or conscripted an array of iron, copper and molybdenum cofactors for the specific enzymes that carry out particular redox transformations. This redox inventory emphasizes both one- and two-electron transfers to and from the nitrogen atoms undergoing chemical change. In particular there are a variety of tetrapyrrolic heme structural variants that serve in novel redox transformations, prefiguring a chapter in Section III on tetrapyrrole macrocycle biosynthesis.

Ammonia (NH_3) with its nucleophilic and basic nitrogen atom is the crossover metabolite between inorganic and organic nitrogen chemical biology. It is the reason nitrogenase was such a valuable invention for life on this planet. The microbes of Section II, elaborating the different parts of the inorganic nitrogen biogeochemical cycle, used that ammonia internally to build the universe of organic nitrogen compounds essential to DNA, RNA and protein construction and function. That capacity also radiated to plants and evolution of multicellular eukaryotes of all types.

We turn to organic nitrogen chemical biology in the nine chapters of Section III.

Inorganic Nitrogen Biochemistry and Dinitrogen Reduction to Ammonia by Nitrogenase

2.1 Introduction to N_2 , NH_3 and NO_3^-

Three molecules dominate the inorganic chemical biology of nitrogen: dinitrogen gas (N_2), ammonia (NH_3) and the deprotonated form of nitric acid, the nitrate anion (NO_3^-) ¹ (Figure 2.1).

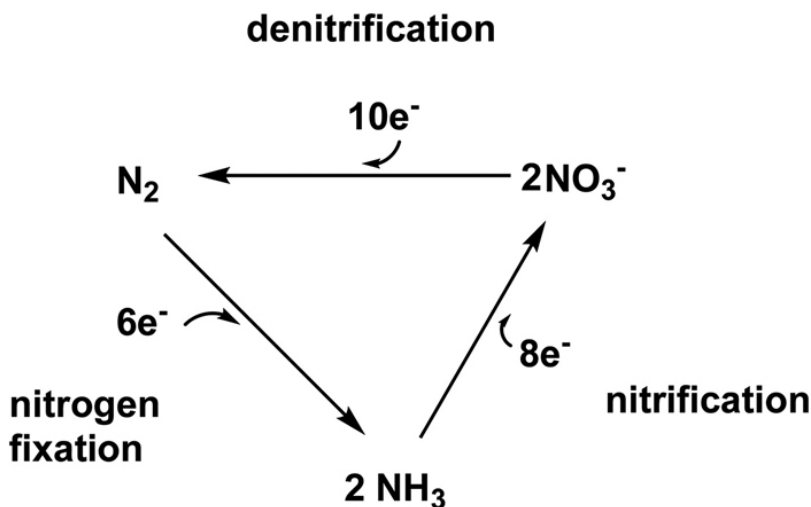


Figure 2.1 Three key molecules in inorganic nitrogen chemical biology: Separated by six electrons (N_2 to NH_3), eight electrons (NH_3 to NO_3^-), and ten electrons (2NO_3^- to N_2). Counting the electrons gives insight into the underlying enzymatic logic.

Of these three, N_2 is by far the most abundant source of nitrogen atoms on the planet. Most familiarly, it is the invisible, tasteless, largely inert gas that comprises 78% of the planetary atmosphere, as noted in Chapter 1. The global atmosphere is

estimated to contain approximately 4×10^{21} g of (di)nitrogen gas. At a molecular weight of 28 g mol^{-1} for N_2 , the atmospheric N_2 content totals about 1.4×10^{20} moles, *none of it in a form usable to organisms*. The lack of chemical reactivity arises from the very strong N–N triple bond, requiring an input of approximately 225 kJ mol^{-1} to break it. In addition to comprising the major volume of air that we breathe in and out, unchanged, N_2 dissolves in water, oceans and freshwater bodies at approximately $20 \text{ }\mu\text{M}$ levels. This dissolved N_2 gas is sufficient to sustain life for those microbes that have evolved nitrogenase enzymes to reduce N_2 down to NH_3 . Indeed, throughout evolution, all life on the planet depended on the ability of those microbes, often in association with plant roots to “fix” nitrogen, meaning to carry out its six-electron reduction to two product molecules of ammonia (NH_3).

The second anchoring molecule of inorganic metabolism is ammonia. Ammonia, NH_3 , like N_2 , is also a gas under physiological conditions (boiling point of $-33 \text{ }^\circ\text{C}$). Highly soluble, up to 305 g L^{-1} at $25 \text{ }^\circ\text{C}$, ammonia can constitute an approximately 18 molar solution at saturation. The pK_a of ammonia is 9.3, the pH at which 50% free ammonia and 50% protonated ammonium ions exist in aqueous solutions.

As the six-electron (and six-proton) reduction product from nitrogenase action by synergistic and free-living species of bacteria, ammonia, largely present as the ammonium ion under physiological conditions, is the first biologically usable form of inorganic nitrogen for metabolism. Ammonia is also the inorganic nitrogen molecule that crosses over to generate all the forms of organic nitrogen compounds in nature (Section III, Chapter 6). As we noted briefly in Chapter 1, ammonia is also generated by bacterial and fungal breakdown of decaying and dead animal and plant biomass through proteolysis and enzymatic deaminations to yield inorganic NH_3 . This pool of ammonia arises mostly from liberation of the amino groups in the amino acid building blocks of proteins.

The third of the major nitrogen forms in inorganic nitrogen chemical biology is the nitrate anion, NO_3^- , the monoanion of nitric acid. HNO_3 , itself a strong mineral acid, a so-called mineral acid because it has no carbon atoms. Nitric acid has a pK_a of -1.3 , so is fully dissociated to nitrate anion and H^+ ions in all biologic milieus. Nitrate is readily diffusible through soils and compared with NH_3 is preferentially taken up by plants that can reduce it back to NH_3 for subsequent organification biochemistry.

This fact previews the cycle of 1 nitrogenase action to “fix” N_2 as NH_3 but then 2 oxidization by eight electrons to produce nitrate as the inorganic nitrogen compound most readily absorbed by plants. Then, 3 the plants have to take it back down to NH_3 to incorporate nitrogen into the amino acids and nucleosides that are the building blocks for proteins and the nucleic acids RNA and DNA (Section III).

2.2 Nitrification, Denitrification and Reductive Assimilation

These three major pathways of inorganic nitrogen chemical biology display nitrogen atoms along the full spectrum of nitrogen redox chemistry. NH_3 with nitrogen in the -3 oxidation state is the most reduced. N_2 is by definition nitrogen with both atoms at oxidation state $\text{N} = 0$. The nitrate anion is the most oxidized with nitrogen at the $+5$ oxidation state (Figure 1.4).

Inorganic nitrogen biology thus operates between $\text{N} = -3$ and $\text{N} = +5$: from NH_3 to NO_3^- , a span of eight electrons. At the outset, this perspective emphasizes that inorganic chemical biology is largely about redox reactions, putting energy in as nitrogen (at $\text{N} = +5$) is reduced, or harvesting energy out as nitrogen is oxidized (from $\text{N} = -3$). The different directionality of the traverses in nitrogen oxidation

states have been accorded distinct nomenclatures. 1

The enzymatic traverse of ammonia to nitrate ions, an energy releasing process, as ammonia is oxidized is termed *nitrification* 2 (making nitrate) and is carried on by certain soil and marine bacteria that make a living from capturing some of the energy released during the inorganic oxidative steps (Figure 2.2A and B). The ability of such microbes (bacteria and archaea) to grow on reduced inorganic substrates (e.g. NH_3) as the sole energy source is also termed lithotrophic or chemolithotrophic growth. In Chapter 3 we will examine how nitrification generates nitrate anions by ammonia oxidation redox enzymology. 3-4 A set of four enzymes convert ammonia to hydroxylamine, nitric oxide and nitrite in that order. Nitrate is then the major source of nitrogen uptake by plant roots.

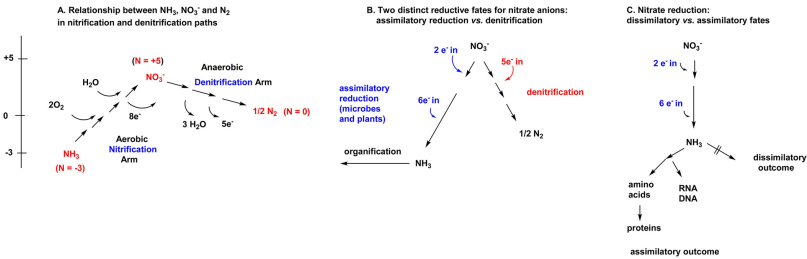


Figure 2.2(A) One way to view the conversion of NH_3 back to N_2 by way of NO_3^- – as an intermediate (nitrification then denitrification). (B) Two distinctive reductive fates for nitrate anions. (C) Assimilatory vs. dissimilatory modes of nitrate reduction.

Microbes that also live among the community of nitrifying bacteria can take the nitrate product and run it the other way. Some of those bacteria are anaerobes. In the absence of O_2 as a possible terminal electron acceptor for metabolite-derived electrons that accumulate as organisms run glycolysis and the tricarboxylate cycle, nitrate reducers can use nitrate as an alternate terminal acceptor. This *could be* termed respiratory denitrification. However, denitrification is a term typically restricted to microbial reductive shunting of nitrate ions to dinitrogen. 5 Therefore, one would categorize the anaerobic bacteria as nitrate respirers. 6 They might also fall under the category of dissimilatory nitrate reducers 7 if the reduced nitrogen products are not being funneled into assimilatory (organification) pathways (Figure 2.2C).

In parallel, plants that have taken up nitrate ions undertake the full eight-electron reduction back to NH_3 which they use as the bridge to the organic nitrogen frameworks of amino acids and nucleobases among other nitrogen metabolites. 8 Both of these eight-electron reduction pathways represent *nitrate reductions*, as a removal of nitrate from the local microenvirons. These are typically termed dissimilatory nitrate reductions when nitrate serves as alternate terminal acceptor for microbes. When plants are reducing nitrate to ammonia for immediate assimilation into the common organic amino acids and nucleosides and nucleotide building blocks of nucleic acids, the process is termed *assimilatory nitrate reduction* (Figure 2.2B and C).

The term *denitrification* in microbiology is usually reserved for microbial reduction of inorganic nitrate ions to a different product, N_2 . 5 Denitrification completes the third arm of the global biogeochemical nitrogen cycle as N_2 is returned, ultimately to the large atmospheric pool (Figures 1.5 and 2.1). Organic nitrogen in the form of decaying, dead biomass is ultimately metabolized to nitrate by soil bacteria and fungi, producing the substrate for denitrification.

2.3 N₂ to NH₃: Biological Fixation of Dinitrogen by Bacterial Nitrogenases

The first step in the nitrogen biogeochemical cycle and the process that provides nitrogen atoms in usable form for organisms to survive and build amino acids, proteins, nucleic acids and other nitrogenous metabolites, is the six-electron reduction of dinitrogen N₂ to two molecules of ammonia, NH₃ ([Figure 2.3](#)).

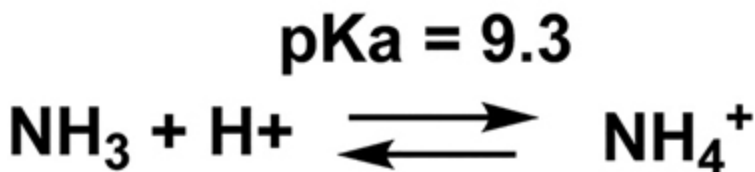
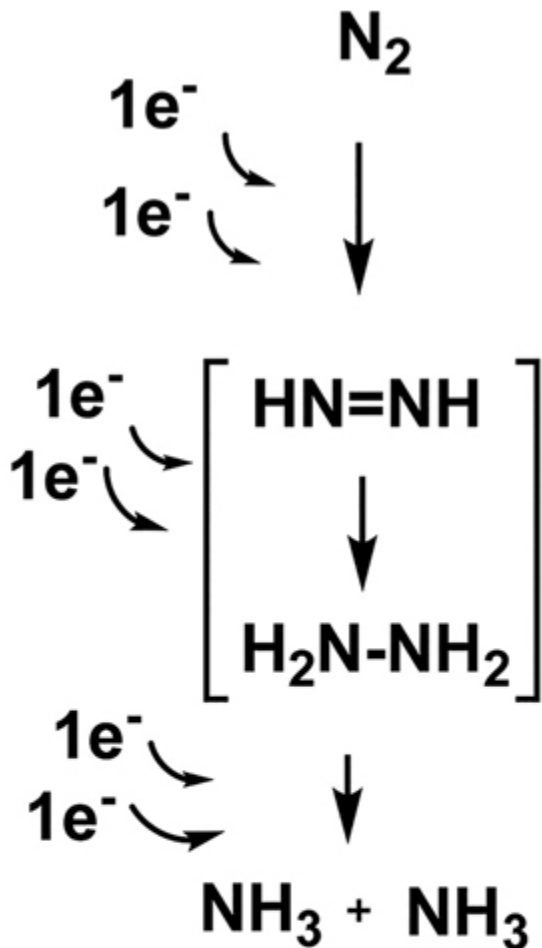


Figure 2.3 Nitrogenase catalyzes six-electron reduction of dinitrogen to two molecules of NH_3 . The two-electron reduced diazene and four-electron reduced hydrazine are proposed to be tightly bound intermediates that do not get released in any catalytic cycle. This is the process of nitrogen fixation.

The global abundance of nitrogen locked up as unreactive N_2 brings to mind the famous line by Coleridge in *The Rime of the Ancient Mariner*: “water, water everywhere, nor any drop to drink”. Bacterial nitrogenases take the highly abundant but essentially chemically inert N_2 and reduce it to the most common form of nitrogen, ammonia, that is incorporated into all the organic scaffolds that enable life (e.g. proteins, nucleic acids, membrane phospholipids). Without doubt, the evolution of nanomachinery to reductively hydrogenate N_2 to two molecules of NH_3 was one of the enabling features of all forms of life on the planet and represents an ancient solution invented in anaerobic atmospheres.

2.3.1 Alternative to Nitrogenase

Throughout the eons of geological time on the earth, up to the beginning of the 20th century, there were two methods of reducing inert atmospheric N_2 to useful, reduced NH_3 . One route was geochemical from lightning discharges that produced enough adsorbed energy in atmospheric N_2 molecules to fragment them, with a variable yield of NH_3 and nitrogen oxides. ⁹ From year to year this can amount to 5–10% of the global amount of N_2 from microbial nitrogenase action.

The second mode is the biological fixation of N_2 to NH_3 by anaerobic or aerotolerant bacteria and blue green algae or cyanobacteria *via* action of the multisubunit enzyme nitrogenase, the subject of the remainder of this chapter.

2.3.2 The Haber–Bosch Industrial Process

Starting in 1908, and through optimized scale up for much of the next 40 years, a third process, the abiotic industrial reduction of N_2 by H_2 at high temperature (approximately 500 °C) and pressure (15–20 atmospheres) over a metal catalyst known as the Haber–Bosch process revolutionized the world by creating millions of tons of NH_3 *abiotically* (Figure 2.4). The unlocked nitrogen reactivity was first used by Germany for explosives but the current worldwide market is for nitrogen from the Haber–Bosch chemistry for nitrates and ureas as fertilizers. ^{10–11} The ability to supply these chemical fertilizers underlies the dramatic increases in crop productivity that have supported the global population expansion from 1 billion people in 1900 to 7 billion in 2020. The contemporary split for NH_3 production globally is close to 50% by nitrogenase biology vs. 50% by the Haber–Bosch industrial process. Both routes accounts for billions of tons per year each, by the parallel biological and abiological processes. An engineering schematic of the industrial scale process is depicted in Figure 2.4.

300 atmospheres and 450 °C

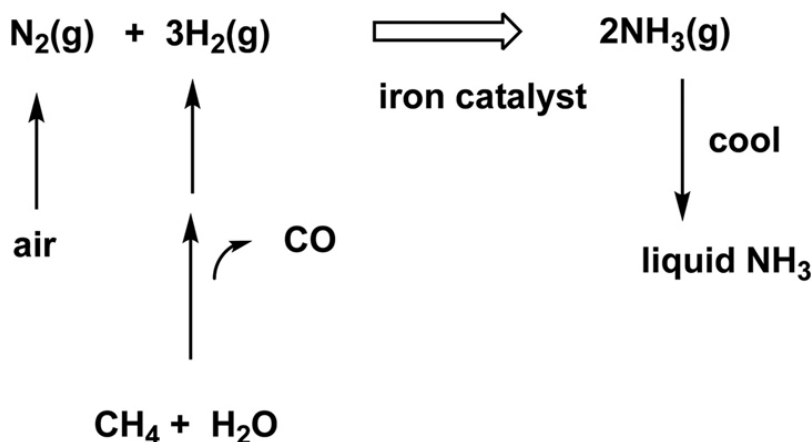


Figure 2.4 The Haber–Bosch process for industrial abiotic hydrogenation of dinitrogen to ammonia over an iron surface feeds half the world through nitrogen fertilizers. The Engineering Schematic Diagram – Reproduced with permission from The Engineering Concepts, <https://www.theengineeringconcepts.com/haber-process-ammonia-manufacturing/>. Reproduced with permission from The NEED Project, www.need.org.

2.4 Nitrogen-fixing Bacteria

The first N_2 -fixing bacteria were isolated by microbiologists at the turn of the 20th century some 120 years ago (see [ref. 12](#)) ([Figure 2.5](#)), although they are thought to have originated about two billion years ago. The nitrogenase nomenclature for the enzyme carrying out N_2 reduction was applied in 1934 by Burk. [13](#) The nitrogen-fixing bacteria are also categorizable as diazotrophs, both strict anaerobes and facultative anaerobes, that can compartmentalize the nitrogenase away from intracellular, deactivating O_2 . Two distinct categories of bacteria contain nitrogenases. Free living clades of bacteria include cyanobacteria ([Figure 2.5A](#)) (*Anabena* and *Nostoc* species), green sulfur bacteria and genera such as *Azotobacter* and *Clostridium*.

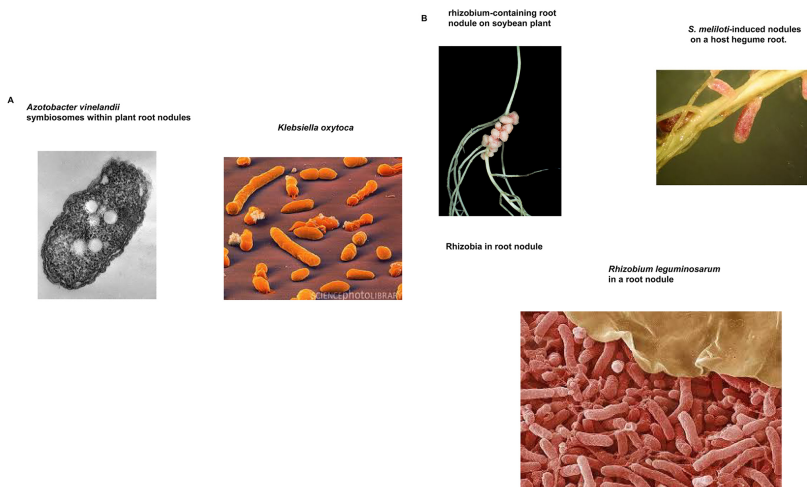


Figure 2.5 Free living or symbiotic nitrogen fixing bacteria: (A) Free living: micrographs of *Azotobacter vinelandii*; Reproduced from ref. 33 with permission from Elsevier, Copyright 2005. *Klebsiella oxytoca*; Image Credit: Eye of Science/ Science Photo Library. (B) *Rhizobium*-containing root nodules on soybean plant. Image Credit: Dr Jeremy Burgess/Science Photo Library. *Sinorhizobium meliloti*-induced nodules on a host legume root. Reproduced from ref. 32, <https://doi.org/10.4137/MBI.S3137>, under the terms of the CC BY 2.0 license <https://creativecommons.org/licenses/by/2.0/>. *Rhizobium leguminosarum* in a root nodule. Image Credit: Power and Syred/Science Photo Library.

Others are mutualists entering into symbiotic interactions with the roots of plants such as legumes: soybeans, clover and peas (Figure 2.5B). These include *Rhizobium*, *Frankia* and *Azospirillum* bacterial species. These mutualists attach to root hairs of plants and use lipo-chito-oligosaccharide fragments termed nodulation factors to induce the plants to fashion nodular organelles. The bacteria enter root cells through induction threads, multiply and morph into bacteroids and become encased in plant-derived symbiosome membranes (Figure 2.5C). The plant–bacterial root nodule symbiosis is one of the most famous and consequential examples of interkingdom mutualisms that benefits both partners. 14:15 The molecular, genetic and metabolite exchange details of nodule formations have been carefully studied and catalogued; the many-faceted regulation and host–donor mutualisms 16:17 are beyond the scope of discussion here.

In these microenvironments the bacteria, specialized species for distinct classes of leguminous plants as examples of specific pairwise symbioses (e.g. *Sinorhizobium meliloti* for legumes from the genera *Medicago*, *Melilotus* and *Trigonella*), turn on the sets of genes to enable reduction of N_2 to NH_3 for short path diffusion to be used by their plant hosts. The energy for the 16 ATPs cleaved per N_2 reduced (*vide infra*) by the bacteria derives from the plant host, largely in the form of malate that then drives oxidative metabolism and ATP synthesis in the bacteroids. Ultimately, the energy for N_2 reduction in these mutualists derives from the ability of plants to turn sunlight into biomass *via* photosynthesis.

2.4.1 Nitrogenase Structures and Reaction Stoichiometries

Purified nitrogenase in active form has been available for some 60 years and subjected to barrages of biophysical, biochemical and genetic approaches to decipher the almost

magical ability of these bacterial enzymes to break the second strongest covalent bond known [carbon monoxide (CO) at $276 \text{ kcal mol}^{-1}$ is stronger than the N–N triple bond at $225 \text{ kcal mol}^{-1}$]. [18:19](#)

We will note below that while molybdenum-containing nitrogenases are the most widespread, bacteria such as *Azotobacter* species will still reduce N_2 even when molybdenum is essentially absent. Instead, they survive by elaborating a separate but homologous vanadium nitrogenase, V-nitrogenase (see [Figure 2.6](#)). This represents a highly unusual and rare use of vanadium in a documented biological context. [20](#) (Another locus of vanadium is in vanadate-dependent haloperoxidases. [21](#)) In the molybdenum-limited environments vanadium ions are substituted, genes are turned on for building a vanadium-containing nitrogenase along with ancillary proteins for assembling the redox cofactors (see next section) and chaperoning them into the inactive apo form of that nitrogenase. Even more rare are “iron only” forms of nitrogenase where an extra iron atom replaces either the molybdenum or vanadium metal when both are absent from the microenvironment. [22](#)

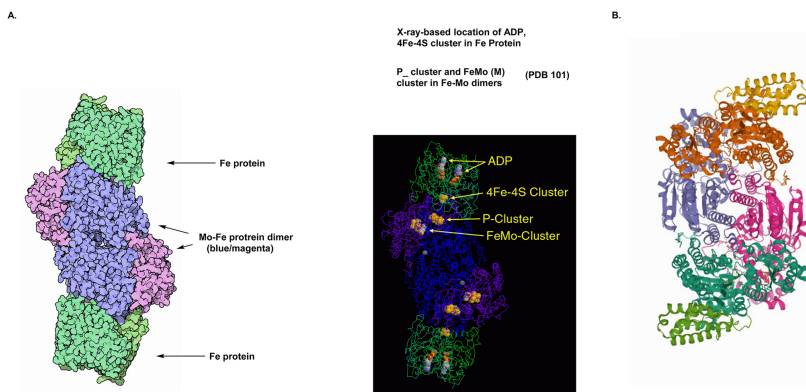


Figure 2.6 (A) The left panel is a space filling model of molybdenum nitrogenase with dimers of FeMo proteins capped at each end by an Fe protein. The right panel shows X-ray density with placement of ADP (from ATP hydrolysis) and the 4Fe–4S cluster in an Fe protein monomer. Each FeMo subunit contains the P cluster and Fe–Mo (M) cluster. Electrons flow in from the 4Fe–4S cluster to P and then M clusters during N_2 reduction. Adapted from PDB101 files. (B) X-ray-determined structure of the vanadium nitrogenase from *Azotobacter vinlandii* grown in Mo-depleted media. Similar overall architecture and cofactor content, with swap of V for Mo, (PDB 5N6Y file).

The most common form, the molybdenum-containing nitrogenases, has two distinct protein components. [18](#) One is a γ_2 dimer containing an $\text{Fe}_4\text{–S}_4$ cluster bridging the subunits and the site for ATP hydrolysis (see [Figure 2.6](#) and discussion) The second component is the FeMo protein, assembled as a pair of $\alpha_2\beta_2$ heterotetramers with the functional units the $\alpha\beta$ -dimers. On the other hand, both the vanadium-containing and iron-only forms of nitrogenases are $\alpha_2\beta_2\delta_2$, oligomers. The γ_2 FeS enzyme feeds electrons into either the FeMo or FeV catalytic component. [20](#) Both the Mo- and V-forms of nitrogenase have been crystallized, the Mo-enzyme many times (including from *Klebsiella oxytoca*) and the V-enzyme most recently in 2017 [23](#) from *Azotobacter vinlandii* ([Figure 2.6](#)).

The stoichiometry of the iron-only nitrogenase is not well established, but [Figure 2.7](#) lists the reaction stoichiometries for the molybdenum and vanadium nitrogenases. Several features are striking. For the more common Mo-enzyme, there is an obligate

liberation of one molecule of molecular hydrogen (gas, H₂) for every molecule of N₂ reduced. ¹⁸ This is addressed in the section on nitrogenase mechanisms below. The reduction of each molecule of dinitrogen requires energy (to break the strong N–N triple bond); experimentally the cleavage of 16 ATPs for each N₂ reduced has been established.

1. Molybdenum-Nitrogenase



2. Vanadium-Hydrogenase

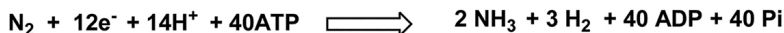


Figure 2.7 Reaction stoichiometries of molybdenum vs. vanadium nitrogenases. Two indicators of less efficient energy coupling in the vanadium nitrogenase are the requirement for 12 electrons to make N₂ with 6 electrons siphoned off to 3 molecules of coproduct H₂. Also, the vanadium enzyme spends 24 more ATPs than the molybdenum enzyme for every molecule of N₂ reduced.

In comparison, the vanadium-nitrogenase is less efficient. First, it leaks electrons into H₂ three times more often than the molybdenum form of the enzyme, diverting 50% of the electron input (6 out of 12 electrons) away from N₂ reduction. In line with the lowered efficiency, 40 ATPs are required for each N₂ reduced. Compared with the 16 : 1 stoichiometry for the Mo-nitrogenase, the V-nitrogenase “wastes” 24 ATPs in each turnover. Put another way the V-nitrogenase is only 32% as efficient as the Mo-nitrogenase, creating some constraints on mechanistic formulations.

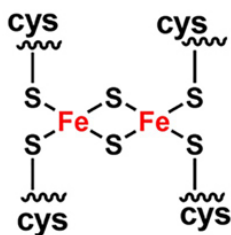
2.4.2 Nitrogenase Redox Cofactor Inventories

2.4.2.1 Iron–Sulfur Clusters

Nitrogenases presumably evolved in diazotrophic bacteria and cyanobacteria two billion years ago when the atmosphere was anoxic. ¹⁸ The enzyme's redox inventory is iron-rich, reflecting the early assembly of inorganic iron–sulfur clusters. ²⁴ Fe–S clusters serve as biological nanomachinery for electron transfers between Fe^{III} and Fe^{II} oxidation states, one-electron transfers. As biological global oxygenation occurred, the reduced forms of iron–sulfur clusters revealed that they are typically oxidized and irreversibly destroyed in the presence of oxygen. Despite this lability, they have survived and persisted as one of the four common assemblages of redox active iron in proteins, with >100 iron–sulfur clusters in humans, presumably in low [oxygen] microenvironments within cells and organelles.

One can find single Fe atoms (e.g. in rubredoxin), Fe₂–S₂ clusters, Fe₄–S₄ cubanes, as presumed precursors to the more complex P and M clusters in nitrogenases (Figures

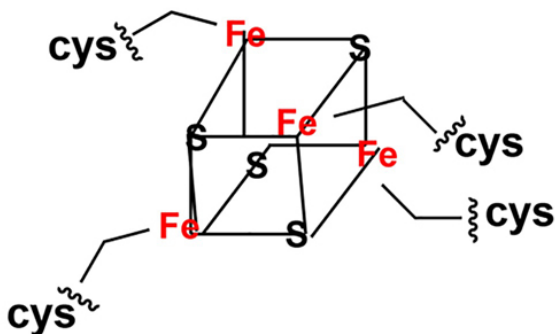
2.8–2.10). 25 The increase in inventory and architectural complexity is presumed to represent successive oligomerizations aided by thiol groups from cysteine side chains as protein ligands and chaperones to stabilize the Fe–S redox elements. We will note below that the P cluster and M cluster in nitrogenase catalytic subunits may have been an early example of optimal biological tuning of biological Fe–S chemistry.



**canonical
 Fe_2S_2 cluster
with 4 Cys
ligands**



**Fe_2S_2 cluster
with 2His, 2 Cys
ligands to iron**



**canonical Fe_4S_4
cluster with 4 Cys
ligands to iron**

Figure 2.8Iron–sulfur clusters of canonical Fe_2S_2 and Fe_4S_4 architectures.

A. Fe_4S_4 bridging to heme
in Sulfite Reductase

B. Nickel-Iron Hydrogenase
Part Structure

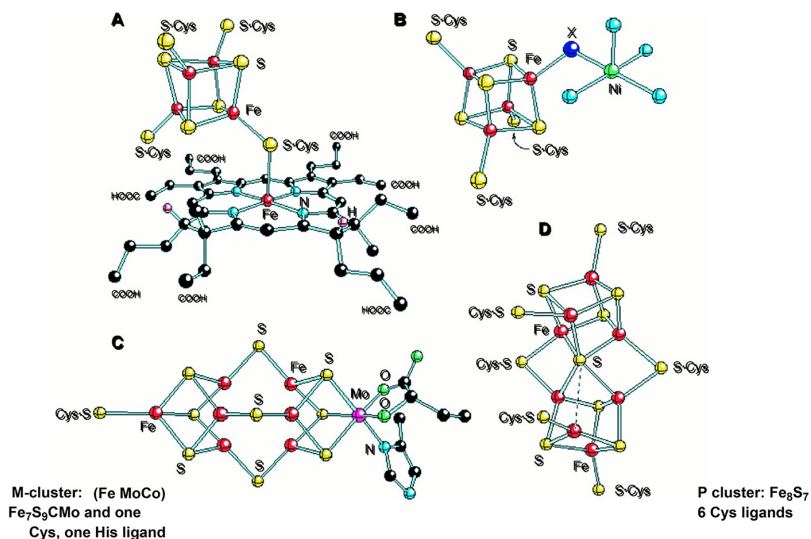


Figure 2.9 Iron-sulfur clusters of four more complex Fe-S architectures including both the Fe_0S_0 P cluster in the Fe-Mo protein and the $\text{Fe}_7\text{S}_9\text{CMo}$ M-type cluster in the Mo-Fe protein subunit of nitrogenase. Reproduced from [ref. 30](#) with permission from The American Association for the Advancement of Science, Copyright 1997.

A. Structure of FeMoCo Core; Central Carbide in Black

B. Structure of the P cluster that feeds electrons
into the MoCo cluster

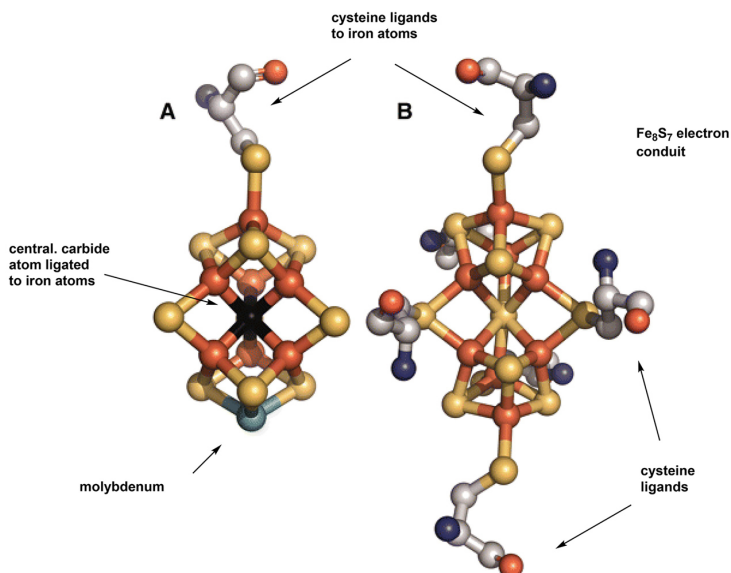


Figure 2.10 Another view of the M cluster that is the active site for N₂ reduction. Reproduced from [ref. 31](#) with permission from The American Association for the Advancement of Science, Copyright 2011.

Iron-sulfur clusters have been studied intensively for more than 50 years, so the primary and secondary literature is immense. Iron-sulfur cluster redox biology permeates almost every aspect of iron redox chemistry, including the electron transfer logic of the ancient enzyme catalysts, the nitrogenases. Interested readers will delve into dedicated iron-sulfur literature. Here, we only deal with a few aspects of cluster architecture and function in nitrogenases. Research is still ongoing into assembly and function of three distinct types of Fe-S architectures therein.

The first is a canonical Fe₄-S₄ cubane type cluster between subunits of the γ_2 dimer of the Fe protein where that cluster serves as the conduit for input of one electron to either the FeMo or FeVa catalytic subunits [26](#) ([Figure 2.9](#)).

2.4.2.2 The P Cluster, M Cluster and V Cluster in Nitrogenases

In the FeMo and FeV catalytic units, the second and third types of Fe-S clusters are found. [26](#) Both are highly complex variants of canonical Fe-S architectures. In the more common Mo-nitrogenase, there is a P cluster (Fe₈S₇) and an M cluster (Fe₇S₉CMo) ([Figure 2.9](#)). The M cluster has also been referred to in the nitrogenase literature as FeMoCo for the FeMo cofactor, but we will tend to use the more recent descriptor, the M cluster.

In the less common V-nitrogenase one also finds the same P cluster (Fe₈S₇) and now a V cluster (Fe₇S₈VC) in place of the M cluster ([Figure 2.11](#)). Studies by Ribbe and colleagues [26](#) at UC Irvine have delineated the cluster assembly logic. Both the P clusters and the M/V clusters appear to originate from Fe₄S₄ cubes that get morphed into Fe₈S₈ double cube architectures during maturation steps. In the P cluster that maturation is an engineered loss of one sulfur atom to go from Fe₈S₈ to Fe₈S₇ ([Figure 2.11](#)). The P cluster is the one-electron functional conduit between the Fe₄S₄ cluster on the Fe protein and the M cluster where the reduction of N₂ takes place ([Figures 2.9 and 2.10](#)).

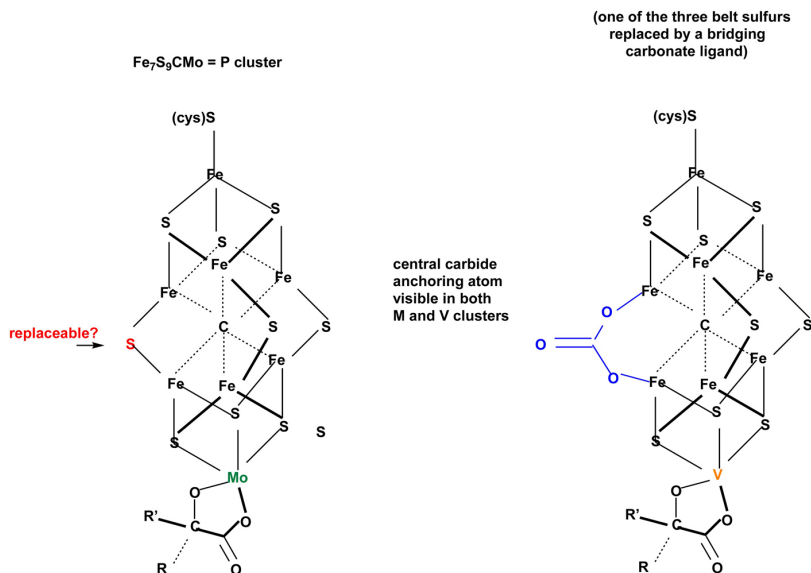


Figure 2.11 Comparison of the corresponding M-type clusters in molybdenum vs. vanadium hydrogenases. Replacement of one of the three belt sulfurs in the Mo cofactor with a bridging carbonate in the X-ray structure of the V nitrogenase.

The M cluster has undergone a remarkable and still incompletely deciphered set of maturations from a presumed Fe_8S_8 precursor. Three immediate items to note in the M cluster (FeMoCo) structure shown in [Figure 2.11](#) are (a) the molybdenum^{VI} atom embedded at one end of the $\text{Fe}_7\text{S}_9\text{MoC}$ scaffold; (b) the homocitrate bidentate ligand to the Mo atom; (c) the stunning finding of a “naked” carbon atom (C = a carbide) as the central anchor to the FeMoCo structure. ²⁶ [In passing, it is known that radical chemistry mediated by *S*-adenosylmethionine (SAM) is involved in carving out the methyl group of SAM as progenitor of the carbide atom in the M cluster. That carbon radical must shed its three hydrogen substituents during a likely multistep insertion and rearrangement of an Fe_8S_8 precursor of the M cluster. ²⁶]

A fourth item is the presence of a ninth sulfur atom, calling attention to the fact that three of the nine sulfur atoms form a belt around the equator of the part of the scaffold nearest the molybdenum atom (see [Figure 2.11](#) for belt sulfurs; and Figure S21 of the original [ref. 27](#)). All four of these maturing modifications are thought to endow the M cluster with the properties it needs to carry out six-electron reduction of N_2 .

The crystal structure of the V-nitrogenase obtained in 2017 ²³ indicated the anticipated high homology of the protein and the redox factor components to the canonical Mo-nitrogenase ([Figure 2.6B](#)). For example, the Fe_4S_4 cubane in the feeder Fe protein and the P cluster (Fe_8S_7) were equivalent. The V cluster, replacing the M cluster, has an altered stoichiometry, at $\text{Fe}_7\text{S}_8\text{VC}$. The missing sulfur atom is one of the three belt sulfurs (S_{3B}) that has been replaced by a bridging carbonate ligand ([Figures 2.6D](#) and [2.11](#) right panel). This substitution offers the possibility that during catalysis the known static snapshots of the M or V clusters at rest may hide a dynamic rearrangement of the belt sulfurs to allow for N_2 binding and reduction. This prospect is taken up in the next section.

2.4.3 Thoughts on Nitrogenase Mechanisms

An enormous compendium of kinetic and spectroscopic observations have been collected over the decades on assembly of the P and M cofactors, on the timing of input of the six electrons sequentially, the hydrolytic expenditure of 16 ATPs, on the spectroscopically detectable states on the M cofactor, on the nature of partially reduced nitrogen intermediates, and many other features (summarized previously 18-27). Some of the mechanistic features of nitrogenase actions have been deciphered, whereas the intimate details of the dynamics of the M cluster and the stepwise reduction of bound N_2 and intermediate states remain cutting edge research topics. 27

Hoffmann *et al.* 18 in a cogent review have summarized many observations into a coherent picture of several aspects of catalysis. One addresses the unusual stoichiometry that eight electrons, not six electrons, are required as input to N_2 for the net six-electron reduction to two NH_3 product molecules. The excess two electrons go into an obligate product molecule of H_2 (Figure 2.12), formed as N_2 binds to the M cluster.

Four electrons into MoFe protein
before N_2 binds: two hydride ion equivalents?

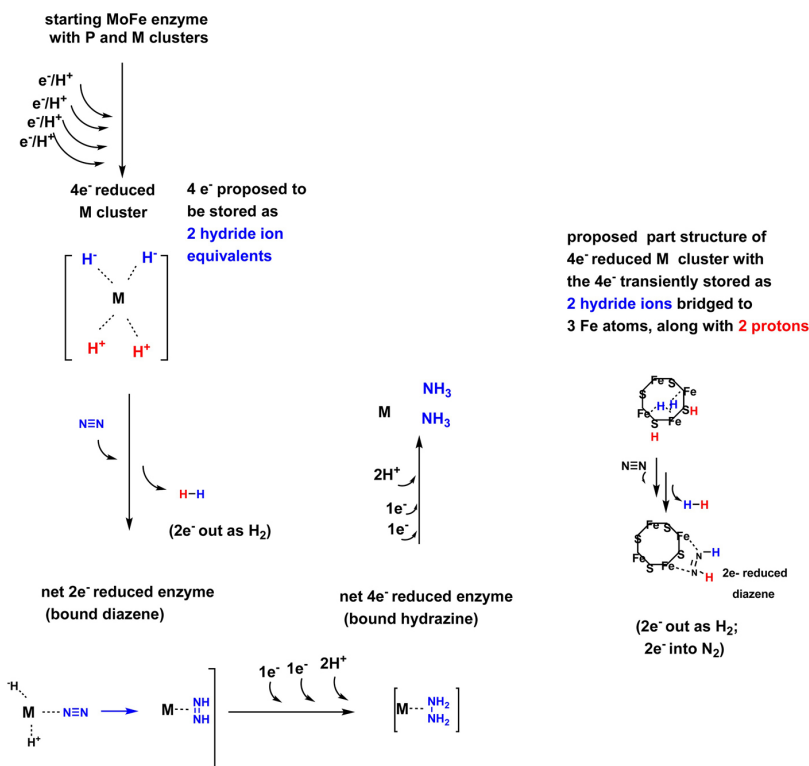


Figure 2.12 Schematic proposal for initial four-electron input into the M cluster before N_2 binds. H_2 is released with two e⁻, while the other 2e⁻ reduce N_2 to bound diazene. Subsequent 2 e⁻ in gives bound hydrazine while the fifth and sixth e⁻ reductively cleave hydrazine to two NH_3 product molecules.

It appears that the first four electrons are stored in the enzyme before N₂ binds productively, with concomitant generation of an H₂ molecule (and thereby using the excess two electrons). Four additional electrons are fed into the M cluster on the way to product ammonia molecules. Although metal nitride intermediates have been considered and are relevant in model chemistry for ruthenium-mediated abiotic reduction, [28](#) it appears that the N–N bond is intact to the very last step in nitrogenase active sites.

2.4.3.1 Bound Hydrazine and Diazine Species

Although firm, direct evidence for diazene (two electron reduced from N₂) and hydrazine (four electron reduced) as “on pathway” intermediates are lacking, the canonical proposals posit that two electrons (of the original four injected by the feeder systems) are utilized for net two electron reduction of M cluster-coordinated dinitrogen to bound diazene (NH=NH). [18](#) In turn, the fifth and sixth electrons would lead to reduction of the sequestered diazene to bound hydrazine (H₂N–NH₂). Finally, the seventh and eighth electrons lead to N–N single bond fragmentation and release of the two NH₃ products ([Figure 2.12](#)), perhaps one ammonia or ammonium ion at a time.

The proposed involvement of dizene and hydrazine are mirrored in reverse order by the four electron-oxidizing hydrazine dehydrogenase in anammox bacteria [29](#) noted in Chapter 3 (see [Figure 3.14](#)). Taken together, nitrogenase and hydrazine dehydrogenase practice nitrogen hydride redox chemistry not seen elsewhere in biological systems.

2.4.3.2 Proposed Storage of Electrons in the M Cluster

Perhaps naively, the presence of the unusual molybdenum atom in FeMoCo, and its substitution by vanadium in some vanadium-containing nitrogenase variants, led some investigators to the initial assumption that the Mo (or V) atom would be redox active and a key focus of the catalytic action. However, the detection of rare iron-only forms of nitrogenases noted above indicates that one or more of the iron atoms in FeMoCo M cluster comprise the catalytic redox center.

Two remaining questions are [1](#) the electron storage capacity of the M cluster (presumably the P cluster is a one electron only redox conduit) and how many electrons need to be stored for electron transfer to ligated N₂ to begin and continue all the way to two NH₃; [2](#) how does N₂ bind productively at the M cluster and undergo controlled reduction?

Hoffmann *et al.* [18](#) proposed that the initial four-electron input (one at a time?) to the M cluster gets stored as two hydride equivalents ($2\text{H}^+ + 4\text{e}^- = 2\text{H}^-$); for example, each hydride would be in bridging configuration between two of the iron atoms in the M cluster. [Figure 2.12](#) shows one such putative bridging hydride (in red) with potential devolution to nonbridging hydride isomeric structures. Insertion of the two hydrides as four electron equivalents between three of the Fe atoms on the Fe_{2,3,6,7} face of the M cluster could form the active species that can bind N₂ productively. This binding event would release H₂ and two of the initially added four electrons, accounting for the obligate formation of one molecule of H₂ as coproduct. It also explains how two of the eight input electrons are diverted away from the six needed for N reduction to two NH₃.

Those remaining two electrons in the partially reduced M cluster (one hydride equivalent, see [Figure 2.13](#)) could be redistributed from the bridging hydride ion to reduce bound N₂ to bound diazene (HN=NH). Repetition of the electron inputs with the fifth and sixth electron would again consume another N–N bond as two more N–H bonds form on the way to bound NH₂–NH₂. The seventh electron sent into the M cluster would result in cleavage of the remaining N–N bond and release of the first

molecule of ammonia with generation of a (delocalized?) Fe-NH_2^- equivalent. The eighth and last electron on could move as a hydrogen atom (H) to yield the second NH_3 , still bound to $\text{Fe}_{\#6}$ and ready to dissociate.

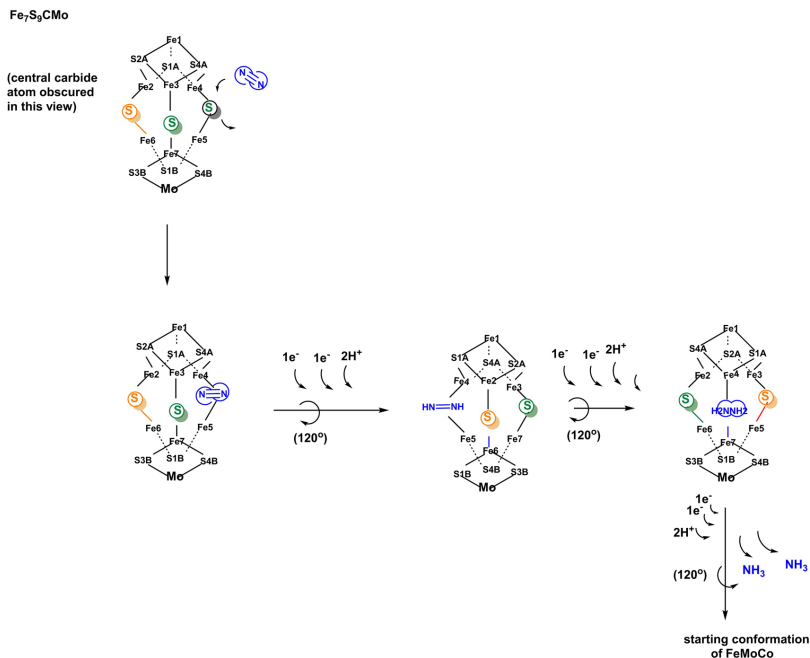


Figure 2.13 Rotation model for binding of N_2 to the M cluster by displacement of one of the three meridian belt sulfur atoms. Each of three $2e^-$ reduction steps would occur with a 120° rotation of the bound nitrogen substrate or intermediates, as suggested by Ribbe and colleagues. ²⁷ It indicates that nitrogenase is a rotary catalyst and the three meridian sulfurs that make up the sulfur belt are key displaceable features of the M cluster in catalytic turnover mode. Adapted from ref. ²⁷ with permission from The American Association for the Advancement of Science, Copyright © 2020 The Authors.

This proposed reaction sequence emphasized that one face of the $\text{Fe}_3\text{S}_4\text{Mo}$ subcluster of the whole M cluster would serve as the electron storage and delivery surface. It also speaks to the observed need for a uniquely complex iron–sulfur architecture. It may be that the corner Mo in this subcluster ($\text{Fe}_3\text{S}_4\text{Mo}$) mediates and optimizes the electronics of the FeSMo cube for the reductive chemistry on N_2 . The use of the $\text{Fe}_3\text{S}_4\text{Mo}$ cube subcluster as a storage device for four electrons, in the form of the two proposed Fe-H-Fe hydride equivalents, would be a remarkable and novel facet of Fe-S cluster redox biochemistry.

2.4.3.3 Nitrogen Swinging through the Belt of the M Cluster?

In addition to the above open question of whether electrons injected into the M cluster from the P cluster are stored transiently as single electrons or as paired hydride ion equivalents is the issue of the structure of the M cluster in its active catalytic mode. The vanadium-nitrogenase cluster has lost one of the three belt sulfurs ²³ (replaced by a bridging carbonate, Figure 2.11) and there is evidence that another of the sulfur

atoms in the M cluster belt can be removed, indicating that the sulfur belt could be dynamic and fluid during dinitrogen reduction cycles.

A recent proposal from Ribbe, Hu, and colleagues [27](#) is the suggestion that N_2 can bind in bidentate mode to two of the M cluster iron atoms, with replacement of one of the three sulfurs of the central trisulfur belt ([Figure 2.13](#) = S21 of [ref. 27](#)). That would be the geometry that leads to productive electron transfer to generate coordinated diazene. They posit that the Mo-Fe half of the M cluster swivels through 120° each time two electrons are or have been delivered ([Figure 2.13](#)). Two more turns around the sulfur belt would first yield the four-electron reduced hydrazine and later the two ammonia product molecules (There is precedent in biology for a rotor enzyme that takes three steps to complete a 360° rotation: the proton-pumping ATP synthase, where electron flow drives the rotations). The completion of the six-electron catalytic cycle would need to see the trisulfur belt reconstituted in the resting M cluster ([Figure 2.13](#)).

2.5 Nitrogen Hydride Biochemistry

Whatever the details of electron and hydrogen transfer, coupled (e.g. two electrons at a time as a hydride ion) or uncoupled (one electron, perhaps as hydrogen atom), nitrogenases catalyze inorganic nitrogen hydride chemistry (diazene, hydrazine and NH_3 ; all have only N-H bonds). These catalysts, invented by nitrogen-fixing microbes billions of years ago with customized FeSCMo/V clusters, are also testament to how difficult it is to cleave reductively that robustly strong triple bond in N_2 and unlock the nucleophilic chemistry available to NH_3 for organification biology. The symbiosis of some forms of rhizobia with legumes globally is the molecular basis for why crop rotation with legumes restores usable forms of nitrogen to nitrogen-depleted crop soils. Estimates of 175 million tons of N_2 reduced by nitrogenase globally each year correspond to approximately 10^{13} moles of N_2 . At 6×10^{23} molecules per mole that is in the range of 10^{36} molecules of N_2 , each reduced yearly to have fed the world through the eons of biological evolution.

References

1. D. Fowler, M. Coyle, U. Skiba, M. A. Sutton, J. N. Cape, R. Reis, L. J. Sheppard, A. Jenkins, B. Grizzetti, J. N. Galloway, P. Vitousek, A. Leach, A. F. Bouwman, K. Butterbach-Bahl, F. Dentener, D. Stevenson, V. Amann and M. Voss The global nitrogen cycle in the twenty-first century, *Philos. Trans. R. Soc. London, Ser. B*, 2013, **368**, 20130164.
2. L. Ackert The role of microbes in agriculture: Sergei Vinogradskii's discovery and investigation of chemosynthesis, 1880-1910, *J. Hist. Biol.*, 2006, **39**, 373-406.
3. R. Hatzenpichler Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea, *Appl. Environ. Microbiol.*, 2012, **78**, 7501-7510.
4. J. P. Zehr and R. Kudela Nitrogen cycle of the open ocean: from genes to ecosystems, *Annu. Rev. Mar. Sci.*, 2011, **3**, 197-225.
5. W. G. Zumft Cell biology and molecular basis of denitrification, *Microbiol. Mol. Biol. Rev.*, 1997, **61**, 533-616.
6. H. K. Marchant, S. Ahmerkamp, G. Lavik, H. E. Tegetmeyer, J. Graf, J. M. Klatta, M. Holtappels, E. Walpersdorf and M. M. M. Kuypers Denitrifying community in coastal sediments performs aerobic and anaerobic respiration simultaneously, *ISME J.*, 2017, **11**, 1799-1812.
7. S. An and W. Gardner Dissimilatory reduction to ammonium (DNRA) as a nitrogen link versus denitrification as a sink in a shallow estuary (Laguna Madre/Baffin

- Bay, Texas), *Mar. Ecol. Prog. Ser.*, 2002, **237**, 41-50.
8. G. Xu, X. Fan and A. J. Miller Plant nitrogen assimilation and use efficiency, *Annu. Rev. Plant Biol.*, 2012, **63**, 153-182.
 9. E. Franzblau and C. Popp Nitrogen oxides produced from lightning, *J. Geophys. Res.*, 1989, **94**, 11089-11104.
 10. V. Smil, *Enriching the Earth; Fritz Haber, Carl Bosch and the Transformation of World Food Production*, MIT Press, Cambridge MA, 2004.
 11. T. Hager, *The Alchemy of Air: A Jewish Genius, a Doomed Tycoon, and the Scientific Discovery That Fed the World but Fueled the Rise of Hitler*, Harmony Books, NY, 2008.
 12. S. E. McGlynn, E. Boyd, J. W. Peters and V. J. Orphan New insights into the evolutionary history of biological nitrogen fixation, *Front. Microbiol.*, 2013, **4**, 201.
 13. D. Burk, *Ergeb. Enzymforsch.*, 1934, **3**, 23.
 14. J. Vasse, B. Camut and G. Truchet Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules, *J. Bacteriol.*, 1990, **172**, 4295-4306.
 15. F. Pini, M. Galaradini, M. Bazzicalupo and A. Mengoni Plant-bacteria association and symbiosis: are there common genetic traits in Alphaproteobacteria?, *Genes*, 2011, **2**, 1017-1032.
 16. T. Coba de la Pena, E. Federova, J. Pueyo and M. M. Lucas The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle, *Front. Plant Sci.*, 2017, **8**, 2229.
 17. V. Oke and S. R. Long Bacteroid formation in the Rhizobium-legume symbiosis, *Curr. Opin. Microbiol.*, 1999, **2**, 641-646.
 18. B. M. Hoffmann, D. Lukoyanov, Z.-Y. Yang, D. R. Dean and L. C. Seefeldt Mechanism of nitrogen fixation by nitrogenase: The next stage, *Chem. Rev.*, 2014, **114**, 4041-4062.
 19. B. M. Hoffmann, D. Lukoyanov, D. R. Dean and L. C. Seefeldt Nitrogenase: a draft mechanism, *Acc. Chem. Res.*, 2013, **46**, 587-595.
 20. Y. Hu, C. Lee and M. W. Ribbe Vanadium nitrogenase: a two hit wonder?, *Dalton Trans.*, 2012, **41**, 1118-1127.
 21. J. M. Winter and B. S. Moore Exploring the chemistry and biology of vanadium-dependent haloperoxidases, *J. Biol. Chem.*, 2009, **284**, 18577-18581.
 22. J. Yang, X. Xie, X. Wang, R. Dixon and Y.-P. Wang Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **E3718-E3725**.
 23. D. Sippel and O. Einsle The structure of vanadium nitrogenase reveals an unusual bridging ligand, *Nat. Chem. Biol.*, 2017, **13**, 956-960.
 24. S. Bandyopadhyay, K. Chandramouli and M. K. Johnson Iron-sulfur cluster biosynthesis, *Biochem. Soc. Trans.*, 2008, **36**, 1112-1119.
 25. Y. Hu and M. W. Ribbe Biosynthesis of the iron-molybdenum cofactor of nitrogenase, *J. Biol. Chem.*, 2013, **288**, 13173-13177.
 26. N. S. Sickerman, M. S. Ribbe and Y. Hu Nitrogenase cofactor assembly: an elemental inventory, *Acc. Chem. Res.*, 2017, **50**, 2834-2841.
 27. W. Kang, C. Lee, A. J. Jasnowski, M. W. Ribbe and Y. Hu Structural evidence for a dynamic metallocofactor during N₂ reduction by Mo-nitrogenase, *Science*, 2020, **368**, 1381-1385.
 28. T. M. Buscagan and D. Rees Rethinking the nitrogenase mechanism: activating the active site, *Joule*, 2019, **3**, 2662-2678.
 29. A. Dietl, C. Ferousi, W. J. Maalkcke, A. Menzel, S. de Vries, J. T. Keltjens, M. S. M. Jetten, B. Kartal and T. R. M. Barends The inner workings of the hydrazine synthase multiprotein complex, *Nature*, 2015, **527**, 394-397.
 30. H. Beinert, R. H. Holm and E. Munck Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures, *Science*, 1997, **277**, 653-659.

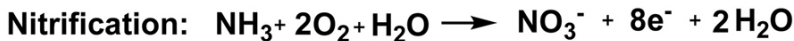
31. K. M. Lancaster *et al.*, X-ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor, *Science*, 2011, **334**, 974-977.
32. R. Rivas, P. Garcia-Fraile and E. Velazquez Taxonomy of Bacteria Nodulating Legumes, *Microbiol. Insights*, 2009, , 51-69.
33. L. J. Liermann, R. L. Guynn, A. Anbar and S. L. Brantley Production of a molybdophore during metal-targeted dissolution of silicates by soil bacteria, *Chem. Geol.*, 2005, **220**, 285-302.

Ammonia Oxidation: Nitrification vs. Anammox

Ammonia oxidizing bacteria in soil are responsible for converting the 120 trillion grams of ammonia and urea fertilizers applied to global soils yearly to more oxidized forms. [1](#) The initial monooxygenation of ammonia to hydroxylamine is the rate-limiting step in the nitrification pathway. [2](#) Some ammonia oxidizers stop after six-electron oxidation to nitrite and provide that molecule to neighboring nitrite oxidizing bacteria to complete the eight-electron oxidation to nitrate ions. [3](#)

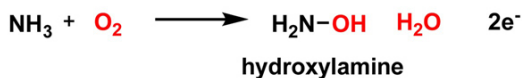
3.1 Nitrification: Ammonia Oxidation to Nitrate

Sergei Winogradsky, a celebrated microbiology pioneer in the late 19th century, observed that bacterial populations can live on NH_3 as a lithotrophic (inorganic) substrate. [4](#) A total of eight electrons are removed as ammonia ($\text{N} = -3$) is oxidized by a set of four enzymes to nitrate ions ($\text{N} = +5$) ([Figure 3.1](#)). Those electrons are fed to electron-transporting cytochromes and ultimately dumped into two O_2 as each is reduced by four electrons to water. The net replacement of three N–H hydrogen bonds in starting substrate ammonia by three N–O bonds in nitrate ions indicates that one or more of the enzymes must carry out N–O bond formations, nitrogen–oxygen chemistry. In turn, it is no surprise that nitrifying bacteria are strict aerobes, although some may be microaerophiles (e.g. microorganism that grow better with 1% O_2 than 20% O_2). Plants take up nitrate ions efficiently from soil, so the ammonia oxidizing bacteria are key providers of the nitrate required by plants for subsequent reductive metabolism and nitrogen organification pathways.

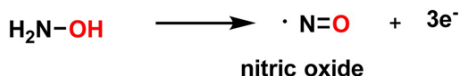


A. Ammonia oxidizing bacteria (*Nitrosomonas*)

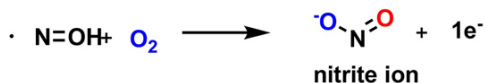
1. Ammonia oxidase



2. Hydroxylamine oxidoreductase



3. Nitric oxide oxidase



B. Nitrite oxidizing bacteria (*Nitrobacter*)

4. Nitrite oxidase

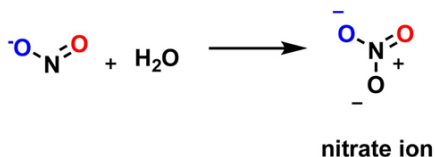


Figure 3.1 Microbial nitrification consists of the 8e^- aerobic oxidation of ammonia by two clades of bacteria. (A) The ammonia oxidizers (e.g. *Nitrosomonas*) convert NH_3 to NO_2^- anion. (B) Then, the nitrite oxidizers (e.g. *Nitrobacter*) remove the last two electrons to produce nitrate anion. The indicated four enzymes act *ad seriatim* in the oxidation pathway with involvement of the nitrogen radical species nitric oxide.

In the event, the first two of the N–O bonds are derived from O_2 and the third N–O bond from water. The use of O_2 for substrate oxygenation by (mono)oxygenase enzymes is typically the province of metalloenzymes on the one hand and vitamin B₂

(riboflavin) type coenzymes on the other hand. ⁵ These two quite distinct strategies (inorganic redox active iron or copper vs. organic flavin heterocycles) represent the two major routes by which organisms have learned to reduce O_2 one electron at a time and carry out transfer of either an $[OH^-]$ or $[OH^+]$ equivalent to substrate.

The four enzymes in ammonia oxidizing bacteria are, respectively, ammonia oxidase, hydroxylamine oxidoreductase, nitric oxide oxidase and nitrite oxidase (Figure 3.1). They are all membrane-affiliated to facilitate electron transfer to respiratory chain quinols and cytochromes. The first six-electron removals generate nitrite anions and are catalyzed by several bacteria including *Nitrosomonas* species. ³ They are known collectively as ammonia oxidizing bacteria.

The last two electron removals, as the nitrite ion is oxidized to a nitrate ion, are typically performed by nitrite oxidizing bacteria (unable to oxidize ammonia). ^{2,3} *Nitrobacter* and *Nitrospira* bacteria (Figure 3.2) often clump together physically with ammonia oxidizing bacteria, consistent with the synergy and intercellular passage by the mixed bacterial populations to take ammonia all the way to nitrate ions. Owners of fish tanks will be familiar with nitrifying bacterial pairs to keep ammonia and nitrite concentrations low in their tanks.

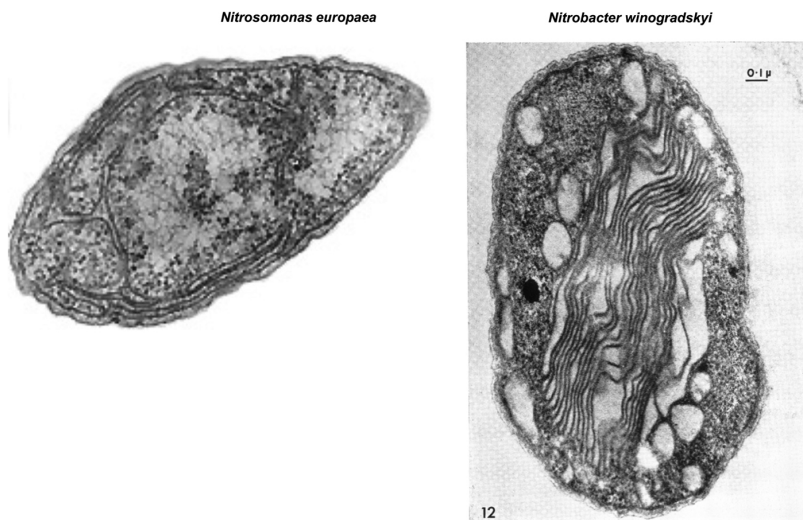


Figure 3.2 Micrographs of the ammonia oxidizing *Nitrosomonas europaea* and of the nitrite ion oxidizing *Nitrobacter winogradskyi*. *Nitrosomonas europaea*: Reproduced with permission from © Woods Hole Oceanographic Institution. *Nitrobacter winogradskyi*: Reproduced from ref. ³⁰ with permission from American Society for Microbiology, Copyright 1965.

Note in Figure 3.1 that although ammonia oxidizing bacteria remove six electrons as NH_3 is oxidized to nitrite ion (NO_2^-), it is not a straightforward removal of two electrons in each of three enzymatic steps. Instead, the pattern is two electrons, three electrons, and then one electron. (We return to one-electron nitrogen chemical biology in Chapter 14 of Section III). The second and third enzymes deserve scrutiny for those unusual reaction stoichiometries. The last step carried out by companion bacteria, the nitrite oxidizers, has a two-electron oxidation stoichiometry.

The four enzymes catalyzing the total of eight electron oxidations involve a broad array of redox active metalloenzyme forms (Figure 3.3), emphasizing the multiple routes that early bacteria had evolved to enable this inorganic nitrogen chemical

biology. 2 These include multicopper centers (ammonia oxidase is a member of the membrane methane monooxygenase family and its two separate copper sites are shown in Figure 3.3A). There is also a 24-heme-containing trimeric enzyme (hydroxylamine oxidoreductase) (Figure 3.3B), and an oxidase that has both nonheme iron and heme iron in its active site (nitric oxide oxidase) (Figure 3.4A shows a proposed bis-iron nitrosyl complex). In the final step, carried out by the nitrite oxidizing bacteria, nitrite oxidase has both iron–sulfur clusters and molybdopterin as redox active metallo-centers (Figure 3.4B).

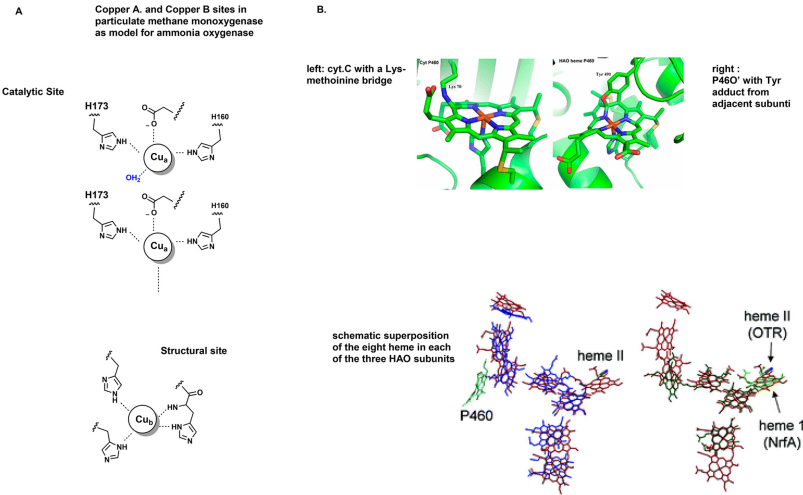


Figure 3.3Redox active cofactors in the first two enzymes of the ammonia oxidizing pathway. (A) Ligands for the coordinatively unsaturated CuA site (probably catalytic metal) and the coordinatively saturated CuB site (probably a structural copper). (B) Two views of the multiple heme tetrapyrrole cofactors in the hydroxylamine oxidoreductase protein. The first view shows five of the eight hemes per subunit. Reproduced from ref. 29 with permission from PNAS. The second view shows the relative orientation of all eight bound heme groups including the single modified P460 heme.

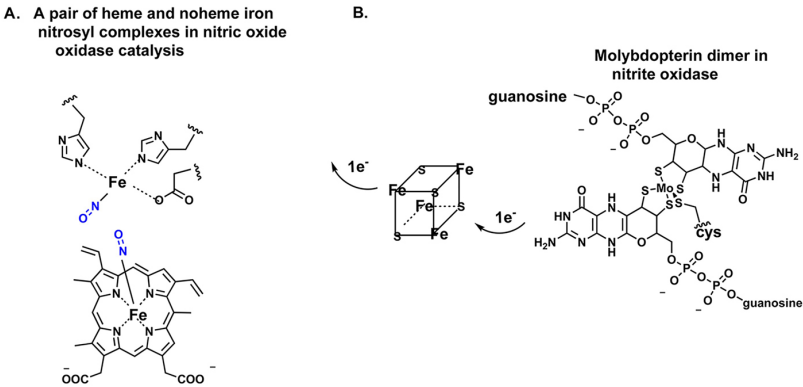


Figure 3.4(A) Heme and nonheme irons in nitric oxide oxidase. (B) Dimeric form

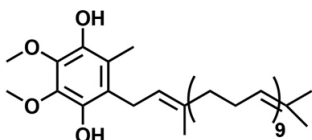
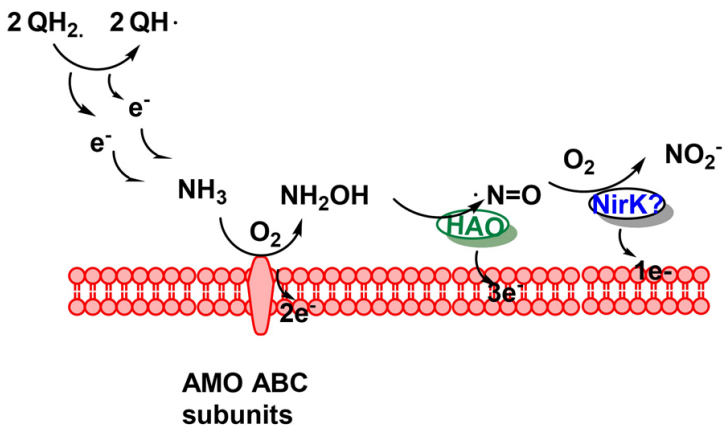
The first enzyme in the pathway, ammonia oxidase, makes an N–O bond using O₂ as ultimate source of oxygen. The third enzyme, nitric oxide oxidase analogously forms a second N–O bond with O₂ as reductive oxygenation agent. The last enzyme, nitrite oxidase probably uses a water molecule as the source of the third of the three N–O bonds in nitrate (see [Figure 3.9](#)).

We take up ammonia oxidizing bacteria first, examining the enzyme inventory and underlying chemical logic for the net six-electron oxidation to nitrite. We then turn to their partners in bacterial consortia to examine the nitrite oxidizer catalytic strategies for the last two-electron step in the nitrification pathway.

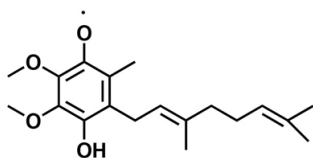
3.1.1 Ammonia Oxidase

The ammonia oxidizing bacteria such as *Nitrosomonas* species can use NH₃ as a sole energy source. [6](#) In this lithotrophic mode they grow slowly, eking out an existence that is limited by the marginally favorable thermodynamics of conversion of NH₃ to NO₂[–] (nitrite). Whole cell studies have been conducted for ammonia oxidation parameters and it has been established that three ammonia oxidase subunits, AmoABC, form membrane-embedded trimers. [2,7](#) The immediate donor of the two electrons, probably one at a time, into the active site copper atoms, required to oxidize NH₃, may be the membrane mobile electron transport carrier coenzyme Q in its reduced QH₂ form ([Figure 3.5](#)).

Proposed Orientation of the Three Enzymes in Ammonia Oxidation to Nitrite



QH_2 = Quinol form of
Coenzyme Q



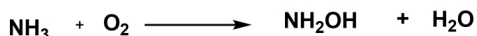
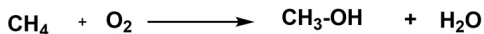
$\text{QH} \cdot$ = semiquinone radical
form of one e^- oxidized QH_2

Figure 3.5 Proposed membrane topology for the three enzymes in the ammonia oxidation pathway on the way to nitrite ions. The net six electrons removed flow into cytochrome acceptors in membrane respiratory chains.

Because ammonia oxidase is an intrinsic membrane protein in bacterial nitrifiers, for many years it has resisted solubilization, purification in an active form and full characterization of its redox center(s) in the active site. The enzyme has high sequence homology to bacterial membrane methane monooxygenase (MMO) and there is crossover reactivity, indicating a similar mechanism and catalytic inventory. Indeed, the isolated MMO enzyme will oxidize ammonia to hydroxylamine as well as its normal conversion of methane to methanol (Figure 3.6). Reciprocally, the *Paracoccus denitrificans* ammonia oxidase will hydroxylate short chain alkanes. ⁷

Both enzymes have a structural and a catalytic copper site

MMO:



Ammonia oxidase

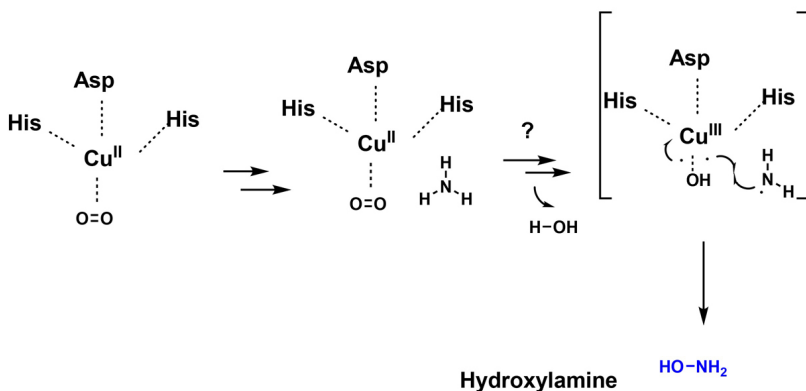
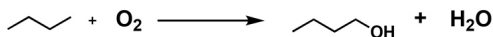
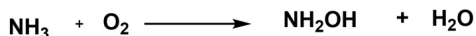


Figure 3.6 Crossover activities between methane monooxygenase and ammonia oxidase: copper-based oxygenases.

Despite earlier claims for binuclear copper in the ammonia oxidase active site, results of later studies [8](#) indicate two functionally discrete copper centers: a structural copper (with four nitrogen ligands) and an active site monomeric copper with two histidine (His) and one aspartate (Asp) side chain ligands ([Figure 3.1A](#)). Water is a dissociable fourth ligand, presumptively replaced by O₂ as oxygen donor to substrate ammonia. For both methane monooxygenase and ammonia oxidase, substrate radical transition states, [CH₃[•], NH₂[•]] captured by an [OH[•]] rebound from a high valent copper-oxide type oxidant, may be the rudiments of the oxygen transfer mechanism ([Figure 3.6](#)). There are several types of precedented copper-based oxygenases, from dopamine beta hydroxylase in mammalian neurotransmitter biosynthesis to copper-containing laccases that create oxidized polymeric products. The bacterial polysaccharide monooxygenase [9](#) has a mononuclear copper site that may be isostructural and isofunctional with ammonia oxidase, illustrating the evolutionary radiation of mononuclear copper oxygenation machinery for different chemical oxygenation tasks.

The oxidase product hydroxylamine is potentially nucleophilic on both nitrogen and oxygen (a so-called alpha effect nucleophile [10](#)) and therefore can be quite toxic to cells by adventitious addition reactions. In fact, it is substantially more nucleophilic

than ammonia and used classically as an analytical tool to capture acyl thioesters as acylhydroxamic acids. ¹¹ Disruption of cellular acyl thioester metabolism would be catastrophic. Thus, hydroxylamine is kept at low concentrations by coupled action of the next enzyme hydroxylamine oxidoreductase.

3.1.2 Hydroxylamine Oxidoreductase

Hydroxylamine oxidoreductase, the second of the four enzymes in the ammonia to nitrate oxidative pathway, is better characterized than ammonia oxidases, ¹² with X-ray structural information on the homotrimeric enzyme ¹³ from *Nitrosomonas europaea* [protein data base (PDB) structure file 1FGJ] (Figure 3.7A). Remarkably, each subunit contains eight molecules of heme iron cofactors. The trimer thus has 24 heme units. Seven of eight of the heme groups in each monomer are of the (conventional) heme c type, in which two cysteinyl residue thiols (Cys229, Cys232) (Figure 3.7B) have added into the vinyl side chains to tether the tetrapyrrolic porphyrins covalently to the protein backbone. However, the eighth heme is modified, differentially reactive and is the locus of NH₂OH binding and oxidation.

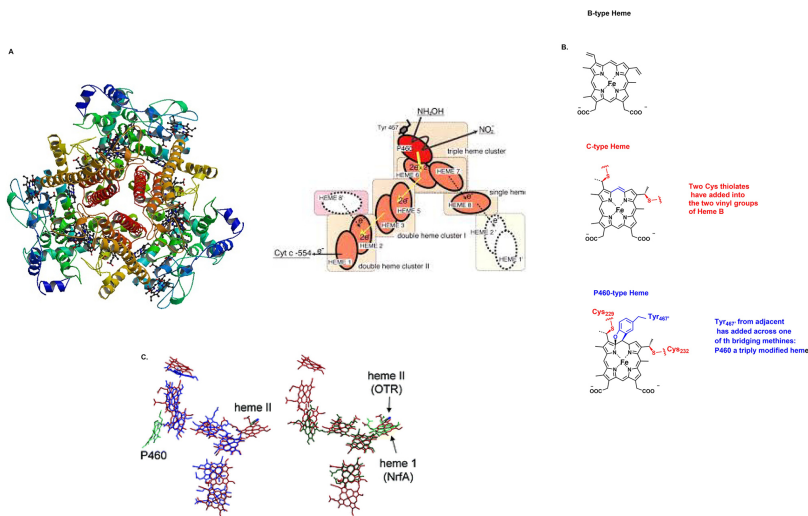


Figure 3.7 (A) X-ray structure of *Nitrosomonas europaea* hydroxylamine oxidoreductase showing a trimer with 24 hemes (PDB1F1J). The right panel shows a schematic for orientation of the eight hemes per enzyme subunit, the oxygenation site on the modified P460 heme and the proposed electron transfer route out to cytochrome c-554 as external one electron acceptor. Reproduced from ref. ¹³ with permission from Springer Nature, Copyright © 1997, Nature Publishing Group. (B) Comparison of the structures of heme-b, heme-c, and the active heme P460 in which Tyr_{460'} from an adjacent subunit has added across a methine bridge of one of the c-type heme groups. Adapted from ref. ¹⁴ with permission from American Chemical Society, Copyright 2013. (C) Reproduced from ref. ²⁸ with permission from the Royal Society of Chemistry.

Elucidation of the structure of this modified heme, termed P460 (for its maximum absorption wavelength at 460 nanometers), required a high resolution X-ray structure of the purified protein. ¹⁴ Figure 3.7B shows that Tyr 467' (prime means from the adjacent subunit), crosslinks the two subunits together through the P460 tetrapyrrolic

chromophore. That crosslink has introduced two new linkages to the tetrapyrrole macrocycle in its adjacent protein subunit. This modification, a probable oxidative self-maturation, disrupts the 20π electron aromaticity of the type c tetrapyrrole macrocycle and distorts or ruffles the plane of the macrocycle. Both these features provide the altered reactivity for electron transfer out of the NH_2OH axially coordinated to the iron atom in the equatorial plane of P460. The two electrons, thought to be transferred out one at a time, will be passed *via* one or more of the attendant c-type hemes out to cytochrome C554 as external protein acceptor (Figure 3.8A).

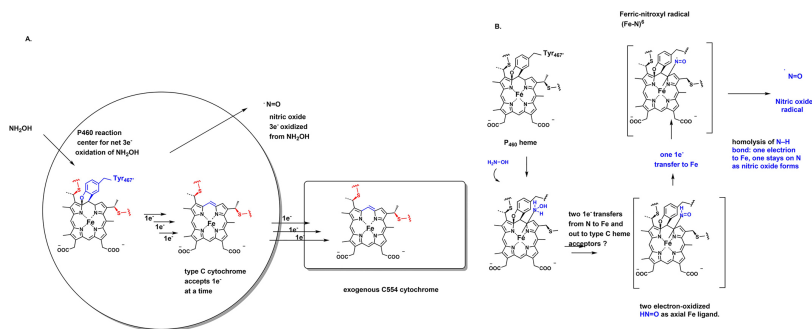


Figure 3.8 (A) Schematic of unusual $3e^-$ oxidation of hydroxylamine to nitric oxide radical (NH_2OH to $[\text{N}=\text{O}]$); (B) proposal for hydroxylamine binding as top axial ligand to Fe in P460 to initiate electron transfer. Two-electron oxidation yields ligated $\text{HN}=\text{O}$. Removal of one more electron yields the nitric oxide radical initially bound as a heme–nitrosyl radical complex before dissociation.

A second unusual feature of hydroxylamine oxidoreductase is its reported ability to catalyze *removal of a third electron* from the resultant bound Fe-HNO intermediate. As suggested in Figure 3.8, one-electron transfer out of the formal $\text{H-N}=\text{O-Fe}$ ligand (perhaps as a hydrogen atom, $\text{H}^\bullet$) would generate the iron-bound nitric oxide. Its release will provide the free radical nitric oxide ($\text{N}=\text{O}$) as the detectable product of hydroxylamine oxidoreductase catalysis. (We will note in Chapters 12 and 14 that nitric oxide is generated in human metabolism as a vasorelaxant signal by human nitric oxide synthase isoenzymes.) The use of 24 heme groups in an enzyme trimer, creating a hydrophobic microenvironment for the hemes and perturbing the ground state redox potentials, to extract three electrons from NH_2OH is a signal example of nitrogen oxidative enzymology. ¹² The array of seven cytochromes c and one P460 heme per subunit create an extended electronic network or battery for electron conduction.

This enzyme is paralleled in its unusual three-electron oxidizing stoichiometry by the mammalian nitric oxide synthases as they convert arginine to nitric oxide and citrulline ¹⁵ (Chapter 12, Figure 12.14) although the redox cofactor nano-machinery (and logic) to remove the first two and then the third electron from the nitrogen atoms are distinct in the two 3-electron enzyme systems.

The $\text{N}=\text{O}$ radical has a high affinity for binding to ferrous iron to form nitroxyl complexes ¹⁶ and that is its signaling fate in mammalian cells. It has a lower affinity for ferric vs. ferrous heme and that may be the mode of dissociation in the hydroxylamine oxidoreductase as proposed in Figure 3.8B.

3.1.3 Nitric Oxide Oxidase

The last step carried out by ammonia to nitrite oxidizing bacteria has a one-electron oxidation stoichiometry. To date the responsible enzyme for the one-electron oxidative removal from the nitric oxide radical to give nitrite has not been identified. Little can be hazarded about cofactor(s) or mechanism(s) other than the assurance that it must be a one-electron transfer metal such as iron or copper.

The combination of a three-electron oxidase for hydroxylamine and then a one-electron removal catalyst for nitric oxide to traverse the net four-electron removal to get to nitrite ion seems a surprising asymmetric division of the redox labor. The reactivity of nitric oxide as a free radical may bias this strategy. ¹⁷ For reasons that are unclear, most ammonia oxidizing bacteria stop at nitrite, six electrons up from ammonia and need specialist bacterial partners to carry out the last two-electron oxidation to nitrate ions. Presumably, the ammonia oxidizers eke out a better living than the nitrite oxidizers in the six to two split in electron removal from the inorganic nitrogen substrates undergoing oxidation.

3.1.4 Nitrite Oxidase (in Nitrite Oxidizing Bacteria)

The last step in ammonia utilization typically requires participation of different clades of bacteria such as *Nitrobacter* or *Nitrospira* species. ¹ The nitrite oxidase is usually a multisubunit membrane-embedded enzyme containing both Fe-S clusters as one-electron conduits and a *molybdopterin cofactor* as the proximal redox center ¹⁸ (Figure 3.4B). *Nitrobacter* nitrite oxidase is in the same family as bacterial dimethyl sulfoxide reductases but runs in the oxidase direction. Up to 30% of *Nitrobacter* membranes can be occupied by these oxidase molecules that are quantitatively important for generation of nitrate ions found in global oceans.

Molybdenum is thus found in two different redox steps in inorganic nitrogen metabolism: in the central FeMoCo Fe₇S₉MoC cofactor of nitrogenase (Figure 2.9) and in the nitrite oxidase at the last step of denitrification (Figure 3.4B). The two environments around the molybdenum atoms are highly distinct. Molybdenum in nitrogenase is embedded at the end of a double iron-sulfur cubane, its role in dinitrogen reduction is still somewhat obscure but not thought to involve redox change in the Mo^{VI} oxidation state. In contrast, the molybdenum atom in nitrite oxidase, and related molybdopterin family oxidoreductases, is coordinated to four thiolate ligands from a dimeric pair of dithiopterins ¹⁹ in a cofactor termed molybdenum guanine dinucleotide (Figures 3.4B and 3.9). Here, the molybdenum atom is the proximal site of electron transfer, undergoing two-electron reduction from the starting Mo^{VI} to the Mo^{IV} oxidation state as a nitrite ion (N = +3) is oxidized by two electrons to a nitrate ion (N = +5).

The intrinsic membrane enzyme nitric oxidase of the nitrifying bacteria has not been extensively studied with respect to mechanism. It can run in the reverse direction, reducing nitrate to nitrite, perhaps foreshadowing the catalytic architecture of the nitrate reductase in denitrifying bacteria (Chapter 4). Also, the sulfite oxidase in sulfite-oxidizing bacteria has a similar molybdopterin guanine dinucleotide cofactor with the bis-thiolenes donating four ligands to Mo^{VI} along with a fifth thiolate from a cysteine residue. The mechanistic proposal for nitrite oxidase in Figure 3.9 is based on the reaction mechanism for sulfite oxidase.

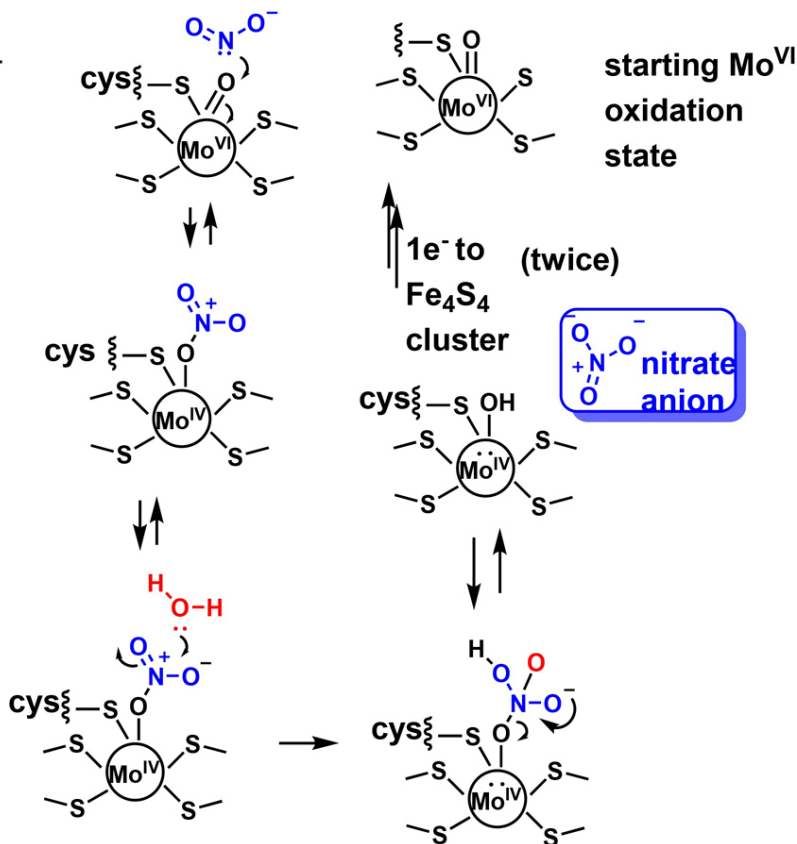


Figure 3.9 Mechanistic proposal for the molybdopterin-heme nitrite ion oxidase, a membrane enzyme catalyzing six-electron reduction without release of any intermediates [the final enzyme in the net $8e^-$ oxidation of ammonia to nitrate (nitrification)].

In the physiological nitrite oxidizing direction, the resting enzyme is thought to have an additional $\text{M}=\text{O}$ ligand that may serve as the electrophile for attack by the nitrite ion oxyanion. That would create a nitrite ester attached to the Mo^{VI} oxidation state. The nitrite portion of that molybdo-nitrite ester adduct could decompose with water attack on nitrogen. Subsequent cleavage of the $\text{Mo}^{\text{IV}}-\text{O}-\text{N}$ bond yields a two-electron reduced Mo^{IV} oxidation state with release of a nitrate anion. The oxidation of nitrite to nitrate is complete, but the electrons are still in the molybdenum atom in the active site. To return the Mo^{IV} back to starting Mo^{VI} for the next catalytic cycle requires reoxidation. Subsequent sequential one-electron transfer from Mo^{IV} to the Fe-S clusters and out to external electron acceptor membrane cytochrome (and ultimately on to O_2) reoxidizes the Mo center from Mo^{IV} to Mo^{V} . Then, by a second electron transfer out, Mo^{V} is converted back to the starting Mo^{VI} oxidation state. ¹⁹

The new N-O bond in nitrate has come from water, not O_2 , as depicted in [Figure 3.9](#). If one looks at the proposed mechanism in the back direction, the enzyme would reduce nitrate ion to nitrite ion, requiring two single-electron inputs from cytochromes or reduced coenzyme QH_2 as possible membrane-based electron donors.

3.1.5 Overview: Nitrification: Ammonia to Nitrate

Oxidative Pathway

The eight-electron oxidative traverse of ammonia to nitrate constitutes the second leg of global biocycling of inorganic nitrogen (Figure 1.5B), where conversion of N_2 to two NH_3 by nitrogenase was the first phase. The ammonia to nitrate metabolic pathway enables lithotrophic *Nitrosomonas* and *Nitrobacter* bacterial species to eke out a living taking the electrons out as N–H bonds are successively replaced by N–O bonds as electrons are passed into membrane electron transport chains on the way to O_2 reduction.

Consistent with other metabolic pathways where O_2 acts as cosubstrate for oxygenases, two redox active metal couples, $\text{Cu}^{\text{I}}\text{--Cu}^{\text{II}}$, $\text{Fe}^{\text{II}}\text{--Fe}^{\text{III}}$, are pressed into service at different stages as electron conduits between O_2 and a nitrogen atom undergoing electron removal. Of the three N–O bonds introduced during transit of an NH_3 to an NO_3^- end product, two of the oxygens are ultimately from O_2 and the last from H_2O as the $\text{Mo}^{\text{VI}}\text{--Mo}^{\text{IV}}$ couple carries the electrons removed (Figure 3.9). Across the ammonia oxidizing bacteria in concert with the nitrate oxidizing bacteria, all eight electrons removed are funneled into membrane electron transport chains to support the anisotropic transmembrane pumping of protons that sets up ATP synthesis as the useful cellular energy currency. In addition to creating the supply of nitrate ions that will be further processed by denitrifying microbes back to N_2 to complete the global nitrogen cycle, nitrate itself is the proximal metabolite taken up by plants for assimilative reductive conversion back to NH_3 for organification (Chapter 5).

3.2 Anammox by the Anaerobic Bacterial Phylum Planctomycetes

By the late twentieth century it had become clear that on the order of 30–50% of the N_2 produced in global oceans could not be accounted for by the known flux of biological denitrification. It turns out that missing percentage arose from actions of anaerobic bacteria combining ammonia with nitrite (Figures 1.5B and 3.10A) to produce the missing N_2 and a molecule of water. ²⁰ This arm of the biogeochemical nitrogen cycle has been termed the anammox arm, for anaerobic ammonia oxidation. While this section could as easily be included in the next chapter on denitrification because the anammox pathway produces N_2 , it is included here on the basis of the oxidation of ammonia. This anaerobic combination of ammonia with nitrite differs fundamentally from the aerobic pathway of ammonia oxidation to nitrate noted above.

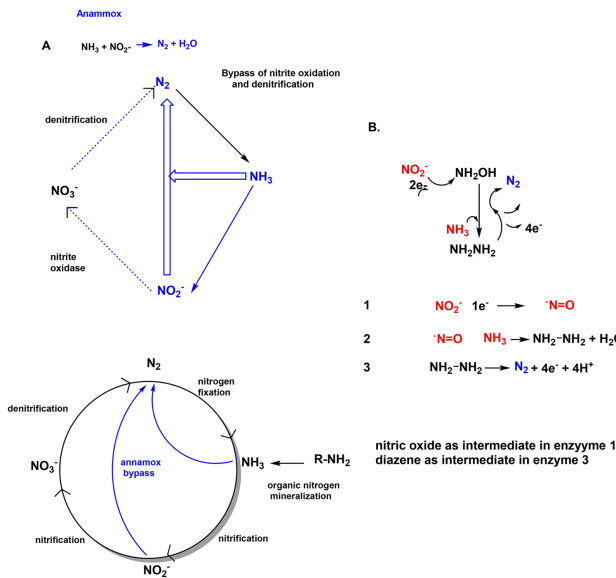


Figure 3.10 The anammox pathway logic. (A) The combination of NH_3 with NO_2^- bypasses both the nitrite oxidation to nitrate and the full denitrification pathway; (B) Schematic of the three-enzyme anammox pathway involving nitrite reductase, hydroxylamine dehydratase and hydrazine dehydrogenase.

The responsible bacteria were isolated only in 1999 from wastewater sludge effluents and found to be members of the bacterial phylum *Planctomycetes*.²⁰ Although five genera from this phylum have subsequently been isolated as anammox utilizers, some 2000 16S rRNA samples have been collected from close relatives, indicating that a broad swath of anaerobes in that phylum can carry out direct conversion of NH_3 to one of the two nitrogen atoms in N_2 . Nitrite ions provide the other nitrogen.

A characteristic of anammox organisms is a large central vacuole where ammonia oxidation occurs. The internal organelle has been termed the anammoxosome.²¹ The doubling time is slow, from one to three weeks, making isolation at scale laborious. The anammox bacteria typically are red in color because of a high content of a cytochrome c isoform that is involved in electron transfer. Despite the slow doubling time, bioengineers at Delft in the Netherlands have run anammox plants at an industrial scale to remove NH_3 from wastewater streams.

The anammox enzymology of this direct conversion of NH_3 and NO_2^- , avoiding the two arms of nitrification then denitrification pathways, is not yet fully elucidated but the strategic outlines are now clear.^{22 · 23} The nitrogen in NH_3 is at oxidation state $\text{N} = -3$. The nitrite oxygen is at $\text{N} = +3$. Each nitrogen in N_2 has the formal oxidation state of $\text{N} = 0$. One could parse the anammox pathway as requiring a three-electron reduction of the nitrite component coupled mechanistically to a three-electron oxidation of the NH_3 component. Along the way three N–N bond forming steps occur to build the strong N–N triple bond in dinitrogen.

The anammox pathway is in fact accomplished by three enzymes acting consecutively as sketched in Figure 3.10B. The first enzyme is a nitrite reductase equivalent to the heme CD enzyme found in denitrifying bacteria²⁴ (to be examined in the next chapter). It carries out single-electron reduction to the nitric oxide radical. That is the freely diffusible substrate for the second enzyme, known as hydrazine synthase. The reaction stoichiometry shown in Figure 3.10B indicates that enzyme

carries out a three-electron reduction as the coupled product hydrazine is produced. ²³ Thus, four electrons have been taken out of the nitrite ion substrate, and the system has switched from N–O to N–H chemistry. Next, the third enzyme hydrazine dehydrogenase ²⁵ catalyzes a four electron *oxidation* of hydrazine to dinitrogen N₂. The three-enzyme pathway is remarkable in putting four electrons in (reduction) and then taking four electrons out (oxidation) to get to N₂ (each N = 0 oxidation state). From just the unusual stoichiometries alone of the nitrite reductase, hydrazine synthase and hydrazine dehydrogenase of 1e[−] and 3e[−] in 4e[−] out some unusual nitrogen chemistry is in play.

We defer discussion of the one-electron transfer catalyst nitrite reductase to the denitrification pathway logic in the next chapter. We will examine the three-electron and four-electron redox logic of the hydrazine synthase and hydrazine dehydrogenase reactions in the remainder of this chapter.

3.2.1 Hydrazine Synthase

The second enzyme of the anammox pathway, hydrazine synthase, is unusual for its product hydrazine, its mechanism and its protein structure. ²² Hydrazine, H₂N–NH₂ is a highly reactive oxidant, used abiotically as rocket fuel. ²⁶ It is highly nucleophilic and toxic when it accumulates in cells. ²⁷ Recall also from Chapter 2 that it has also been postulated to be the tightly held intermediate four electrons down from substrate N₂ in nitrogenase catalytic cycles. In those nitrogen-fixing bacteria free hydrazine is not detected during N₂ reduction. This enzyme is unique in converting an N–O bond to an N–N bond as its NH₂OH substrate gets transformed into NH₂NH₂.

The anammox hydrazine synthase is an intrinsic membrane protein in the organellar anammoxosomes and functions as a dimer of heterotrimers. Each heterotrimeric component has an α, β and γ subunit. ²² The X-ray structure of the *Kuenenia stutzgartii* enzyme, a dimer of trimers, is shown in part in [Figure 3.11](#) with schematic placement of multiple heme cofactors and a tightly bound zinc ion. There are two separate heme subsites, one binding nitrite and feeding in three sequential electrons from the highly abundant, reduced cytochrome c electron donor membrane proteins. The second subsite is accessed by two tunnels. One leads to the outside and is the tunnel by which NH₃ enters to that N–N bond-forming heme site. The second tunnel is internal, leading from the nitrite reduction site to the N–N bond-forming heme subsite ([Figure 3.12](#)). The intermediate diffusing through that internal tunnel is nascent hydroxylamine, the product from the nitrite reduction subsite. It becomes the second substrate (NH₃ is the first) at the second subsite.

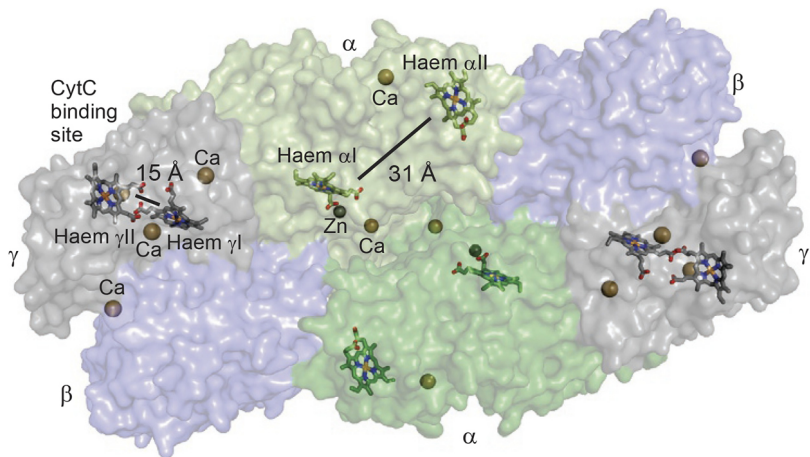


Figure 3.11 Multiple heme cofactor tetrapyrroles in hydrazine synthase, the second enzyme in the anammox pathway. Reproduced from [ref. 22](#) with permission from Springer Nature, Copyright 2015.

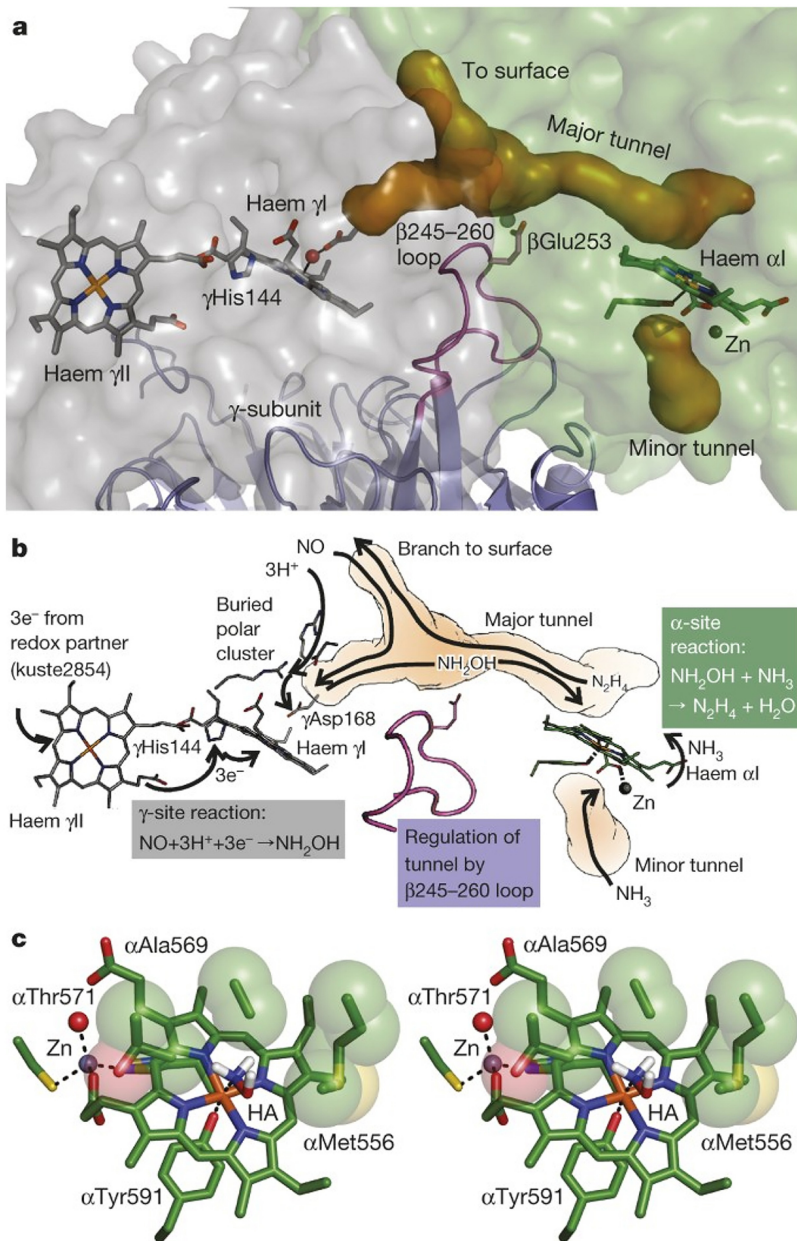


Figure 3.12 Two separate subsites, one for three-electron reduction of nitric oxide to hydroxylamine at heme- γ the other for hydroxylamine displacement by the second substrate ammonia to generate hydrazine. The nascent hydroxylamine diffuses from left to right in parts a and b to get to the ammonia binding site. Two diffusion tunnels for hydroxylamine and for ammonia are noted. Reproduced from [ref. 22](#) with permission from Springer Nature, Copyright 2015.

The three-electron reduction of nitrite to NH_2OH occurs in the central γ -subunit. The N–N bond coupling between NH_2OH and NH_3 occurs in the peripheral α -subunit

(Figure 3.13). In the central γ -subunit are two heme cofactors close to each other. One is a c-type heme with protein histidine residues as both axial ligands (heme γ_1). The other heme (γ_2) appears by X-ray analysis to have a third Cys residue added in to one of the methine bridges of the heme, creating an altered geometry and redox potential. Also, heme- γ_2 has a displaceable water molecule as the top axial ligand and is the site of binding of nitric oxide. Three electrons are added, presumably one at a time, to the nitrite-heme- γ_2 adduct, along with proton transfers: first to yield an axially ligated H-N=O (not shown) and then hydroxylamine (NH_2OH) as a nascent three-electron reduced species (Figure 3.12).

The NH_2OH does not escape to the medium from the enzyme γ -subunit but instead is funneled through the internal protein tunnel connecting the γ - and α -subunits, reaching and coordinating to heme α -1 (Figure 3.12). That internal tunnel is labeled the major tunnel in Figure 3.12. A second minor tunnel from the outside to that second subsite is the entryway for NH_3 to gain entrance.

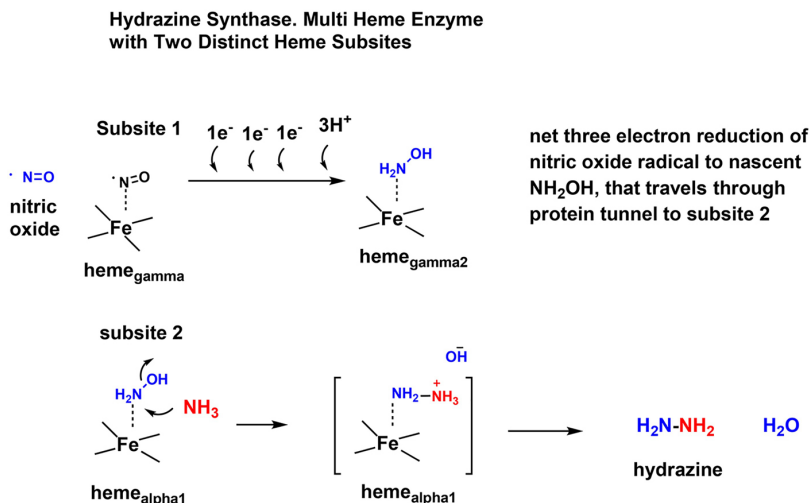


Figure 3.13 Schematic of N-N bond-forming step. Reproduced from ref. 22 with permission from Springer Nature, Copyright 2015.

In this second active site NH_3 binds and is thought to act as a nitrogen nucleophile on the heme- α 1-coordinated NH_2OH . This is the N-N bond-forming step, along with cleavage of the N-O bond to result in liganded hydrazine (Figure 3.13). The released products are a water molecule and hydrazine, $\text{H}_2\text{N-NH}_2$ (both nitrogens are at oxidation state $\text{N} = -2$).

This three-electron *in situ* reduction, the constrained diffusion of nascent NH_2OH to the adjacent α -subunit for capture by a different heme group, followed by N-N bond coupling between exogenous NH_3 and the axially ligated NH_2OH , is all together a remarkable piece of inorganic nitrogen chemistry. The virtuosity of heme is on full display. Dietl and coworkers 22 noted the analogy between the enzymatic route to hydrazine from NH_2OH and NH_3 and the industrial process where one molecule of NH_3 is converted to chloramine (NH_2Cl) and then functions as the electrophile on attack by a second molecule of NH_3 to yield hydrazine and a chloride ion. Getting hydroxylamine to react as an electrophile, rather than the nucleophile that it is at both oxygen and nitrogen when free in solution, is a remarkable reversal of polarity and is ascribed to coordination of the nascent hydroxylamine to the heme iron atom,

facilitating breakage of the N–O bond as NH_3 attacks.

3.2.2 Hydrazine Dehydrogenase

The third of the three enzymes in the anammox pathway is a hydrazine dehydrogenase presumed to oxidize the electron-rich hydrazine to a diazene intermediate and then on to N_2 as released product **25** (Figure 3.14A). The four electrons removed then exit one at a time to the abundant cytochrome c protein molecules (Fe^{3+} acceptor state) in the anammoxosome membrane. Thus, in contrast to enzymes 1 and 2 that perform reductions (of 1e^- and 3e^- , respectively), enzyme 3 does the reverse, it *oxidizes rather than reduces* its substrate by four electrons. The net input of electrons to the anammox system is zero, while a $\text{N} = +3$ (nitrite) molecule and an $\text{N} = -3$ (ammonia) molecule have been elegantly coupled to form the immensely strong N–N triple bond in N_2 .

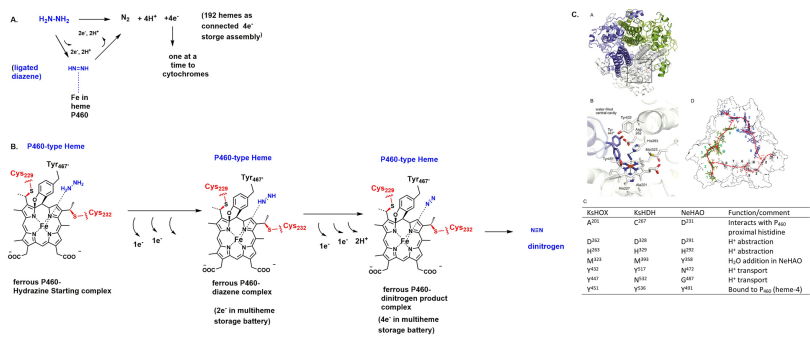


Figure 3.14 Hydrazine dehydrogenase removes four electrons as NH_2NH_2 is doubly desaturated to N_2 . (A) Enzyme reaction stoichiometry. (B) Each of the 24 subunits contains eight hemes (for a total of 192), one of which is modified as heme P460 from a tyrosine addition onto the heme from an adjacent subunit. (C). Proposed mechanism for removal of four electrons from hydrazine axially coordinated to P460. Reproduced from ref. 25 with permission from American Society for Biochemistry Molecular Biology, Copyright 2016.

Hydrazine dehydrogenase, has also been purified and characterized from *Kuenenia stuttgartiensis*, based originally on its homology to the octaheme hydroxylamine oxidoreductases (HAO) noted above in the *aerobic* bacterial nitrification pathways (NH_3 to NO_3^-) (Figure 3.7). The commonality of the multiheme array architecture indicates underlying unifying logic among the hydroxylamine oxidoreductase and the hydrazine dehydrogenase. **25** Both NH_2OH and NH_2NH_2 are electron-rich highly nucleophilic molecules but hydrazine may be in protonic equilibrium with the $\text{H}_2\text{N-NH}_3^+$ cation at pH 6 where the hydrazine dehydrogenase operates in the anammox bacteria.

Both the hydroxylamine oxidoreductase of the ammonia oxidation pathway (Figure 3.7) and the hydrazine dehydrogenase of the anammox bacteria are composed of octaheme-containing protein subunits, linked together by the covalent capture of one of the eight hemes by a tyrosyl residue from the neighboring subunit (Figure 3.14B). Thus, the minimal unit is a crosslinked trimer with $8 \times 3 = 24$ heme groups (21 as cytochrome c, and three as heme P460). These trimers associate to octamers, containing 24 subunits total with an almost astronomical inventory of 192 hemes (!), 24 of which are the P460 modified hemes that are the sites of catalytic action. The

molecular weight of each hydrazine dehydrogenase 24mer is 1.7 million Daltons. Even though the ancillary hemes act as temporary electron storage units, this seems like a profligate expenditure of almost 200 hemes in each hydrazine dehydrogenase complex.

Remarkably, the *K. stuttgartensis* enzyme binds the metabolite hydroxylamine (*vide supra*) but does not oxidize it. Instead, the dehydrogenase is highly specific for hydrazine as oxidizable substrate, converting it all the way to N_2 as released product [25](#) ([Figure 3.14C](#)). The iron atom at the center of the P460 modified tetrapyrrole in each of the eight catalytic subcomplexes binds three ligands: the starting NH_2-NH_2 , intermediate diazene ($HN=NH$), and product N_2 as top axial ligands before N_2 release ([Figure 3.14C](#)). The presumption is that the four-electron oxidation proceeds in four individual steps, with the 168 c-type heme interconnected electronic network serving as a multiheme storage battery for those four electron stages. At some point each of those four stored electrons in the internal heme network are passed, one at a time, out to external cytochrome membrane proteins. The hydrazine and diazene usages are reminiscent of the nitrogenase reaction, but run in reverse.

More of the strategies and mechanisms in anammox microbiology and chemistry remain to be deciphered, in this most recent of the five arms of inorganic nitrogen global cycling ([Figure 1.5](#)). At this juncture we can note that in all three enzymes ([Figure 3.9B](#)), the nitrite to nitric oxide reductase, the nitric oxide to hydrazine synthase and the hydrazine to N_2 double dehydrogenase, nature has called in each instance on heme iron chemistry. Heme variants are key actors in all three transformations; the hydrazine dehydrogenase octamer of trimers with 192 heme groups as a phalanx of electron transfer armamentarium is a stunning marshalling of redox artillery ([Figure 3.14B](#)).

In both the nitrification pathway (ammonia to nitrate) and then the anammox pathway just described, the bacteria switch between nitrogen hydride and nitrogen oxide chemistry. In ammonia oxidation, two of the oxygens in product nitrate ion come from O_2 , the third from water. In the anaerobic anammox pathway O_2 is avoided altogether. The conversion of a nitrite ion [23](#) to one half of an N_2 molecule is a net three-electron reduction, while NH_3 undergoes a comparable three-electron reduction, ending up with N_2 , lacking both hydrogen and oxygen substituents.

3.3 Ammonification: Degradative Conversion of Organic Nitrogen Metabolites to NH_3

The fifth contributor to the global biogeochemical cycling of inorganic forms of nitrogen includes all the degradative pathways that lead from organic nitrogen metabolites to liberate NH_3 . This is the domain of soil fungi and bacteria that act on decaying plant and animal biomass. This involves hydrolysis of the major biopolymers, nucleic acids, proteins, polysaccharides and the phospholipids that make up animal and plant cell membranes back to monomeric building blocks.

The two biopolymers replete with organic nitrogen are the nucleic acids, both RNA and DNA, and all the proteins of the cell ([Figure 1.1](#)). A third group would be the choline, ethanolamine, and serine head groups in membrane phospholipids, in their many millions in every cell. Ammonification then represents the liberation of NH_3 from the amino acid fragments of proteins and the purine and pyrimidine base fragments of nucleic acids. Also included would be deamination of the phospholipid head group fragments and hydrolysis of ureas. These processes are sometimes termed mineralization of nitrogen.

The kinetics of organic nitrogen mineralization will vary depending on terrain, climate and microbe composition. The halftime for turnover of organic nitrogen to

ammonia in a Virginia salt marsh was observed to range from 5 to 14 days, indicating the robust activity of soil microbes, fungi included. At the enzyme level, responsible catalysts include amino acid oxidases, amino acid ammonia lyases, dehydrogenases, peptidases, ureases, amidases and others. They are examined in the context of organic nitrogen substrates in Section III. In addition to release of ammonia some organic nitrogens are oxidized all the way to nitrate ions in aerobic degradative pathways.

References

1. M. M. M. Kuypers, H. Marchant and B. Kartal The microbial nitrogen-cycling network, *Nat. Rev. Microbiol.*, 2011, **1**, 1-14.
2. L. E. Lehtovirta-Morley Ammonia oxidation: ecology, physiology, biochemistry and why the must all come together, *FEMS Microbiol. Lett.*, 2018, **365** 10.1093/femsle/fny058.
3. L. W. Belser Population ecology of nitrifying bacteria, *Annu. Rev. Microbiol.*, 1979, **33**, 309-333.
4. S. Winogradsky Recherches sur les organismes de la nitrification, *Ann. Inst. Pasteur*, 1890, **4**, 213-221.
5. C. T. Walsh and Y. Tang, *Biosynthesis of Natural Products*, Royal Society of Chemistry, London, 2018.
6. S. Gilch, O. Meyer and I. Schmidt A soluble form of ammonia monooxygenase in *Nitrosomonas europaea*, *Biol. Chem.*, 2009, **390**, 863-873.
7. J. W. Moir, L. C. Crossman, S. Spiro and D. J. Richardson The purification of ammonia monooxygenase from *Paracoccus denitrificans*, *FEBS Lett.*, 1996, **387**, 71-74.
8. M. Ross, F. MacMillan, J. Wang, A. Nisthal, T. J. Lawton, B. D. Olafson, S. L. Mayo, A. C. Rosenzweig and B. M. Hoffmann Particulate methane monooxygenase contains only mononuclear copper centers, *Science*, 2019, **364**, 566-570.
9. G. R. Hemsworth, E. Taylor, R. Q. Kim, R. C. Gregory, S. J. Lewis, J. P. Turkenburg, A. Parkin, G. J. Davies and P. W. Walton The copper active site of CBM33 Polysaccharide Oxygenases, *J. Am. Chem. Soc.*, 2013, **135**, 6069-6077.
10. W. P. Jencks and J. Carrioulo Reactivity of nucleophilic reagents toward esters, *J. Am. Chem. Soc.*, 1960, **82**, 1778-1786.
11. D. L. Dunjlemna, Y. Hirata, K. A. Totaro, D. T. Cohgen, C. Zhange, Z. P. Gates and B. L. Pentelute Amide-forming chemical ligation via O-acyl hydroxamic acids, *Proc. Natl. Acad. Sci. U.S.A.*, 2018, **115**, 3752-3757.
12. J. D. Caranto and K. Lancaster Nitric oxide is an obligate nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc. Natl. Acad. Sci. U.S.A.*, 2017, **114**, 817-822.
13. N. Igarashi, H. Moriyama, T. Fujiwara, Y. Fukumori and N. Tanaka The 2.8 angstrom structure of a hydroxylamine oxidoreductase from a nitrifying chemoautotrophic bacterium *Nitrosomonas europaea*, *Nat. Struct. Biol.*, 1997, **4**, 276-284.
14. P. Cedervall, A. Hooper and C. M. Wilmot Structural studies of hydroxylamine oxidoreductase reveal a unique heme cofactor and a previously unidentified interaction partner, *Biochemistry*, 2013, **52**, 6211-6218.
15. J. T. Groves and C. C. Wang Nitric oxide synthase: mechanisms and models, *Curr. Opin. Chem. Biol.*, 2000, **4**, 686-695.
16. W. Flores-Santana, D. Salmon, S. Donzelli, C. H. Switzer, D. B. Sudhar, L. Ridnour, R. Cheng, S. A. Glynn, N. Paolocci, J. M. Fukuto, K. M. Maranda and D. A. Wink The specificity of nitroxyl chemistry is unique among nitrogen oxides in biological systems, *Antioxid. Redox Signal.*, 2011, **14**, 1659-1674.
17. D. D. Thomas, L. Ridnour, J. S. Isenberg, W. Flores-Santana, C. H. Switzer, S.

- Donzellie, P. Hussain, C. Vecoli, N. Paolocci, S. Ambs, C. Colton, C. Harris, D. D. Roberts and D. A. Wink The chemical biology of nitric oxide. Implications in cellular signaling, *Free Radical Biol. Med.*, 2008, **45**, 18-31.
18. T. Yamanaka and Y. Fukumori The nitrite oxidizing system of *Nitrobacter winogradskyi*, *FEMS Microbiol. Rev.*, 1988, **4**, 259-270.
19. J. L. Johnson, N. Bastian and K. V. Rajagopalan Molybdopterin guanine dinucleotide: a modified form of molybdopterin identified in the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides forma specialis denitrificans*, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 3190-3194.
20. M. Strous, E. Pelletier, S. Mangenot, T. Rattei, A. Lehner, M. W. Taylor, M. Horn, H. Daims, D. Bartol-Mavel, P. Wincker, V. Barbe, N. Fonknechten, D. Vallenet, B. Sgurens, C. Scheniwitz-Truong, C. Medgue, A. Coloongro, B. Snel, B. E. Dutilh, H. J. M. O. den Camp, C. van der Drift, I. Cirpus, K. T. de Pas-Schoonen, H. R. Harhangi, L. van Niftrik, M. Schmid, J. Keltjens, J. van de Vosseberg, B. Kartal, H. Meier, D. Frishman, M. A. Huynen, H.-W. Mewes, J. Weissenbach, M. S. M. Jetten, M. Wagner and D. Le Pasluier Deciphering the evolution and metabolism of an anammox bacterium from a community genome, *Nat. Rev. Immunol.*, 2006, **440**, 790-794.
21. L. A. van Niftrik, J. Fuerst, J. S. S. Damste, J. G. Kuenen, M. S. M. Jetten and M. Strous The anammoxosome: an intracytoplasmic compartment in anammox bacteria, *FEMS Microbiol. Lett.*, 2004, **233**, 7-13.
22. A. Dietl, C. Ferosi, W. J. Maalcke, A. Menzel, S. de Vries, J. T. Keltjens, M. S. M. Jetten, B. Kartal and T. R. M. Barends The inner workings of the hydrazine synthase multiprotein complex, *Nature*, 2015, **527**, 394-397.
23. L. B. Maia and J. Moura How biology handles nitrite, *Chem. Rev.*, 2014, **114**, 5273-5357.
24. C. Sparacino-Watkins, J. Stolz and P. Basu Nitrate and periplasmic nitrite reductases, *Chem. Soc. Rev.*, 2014, **43**, 676-706.
25. W. J. Maalcke, J. Reimann, S. de Vries, J. Butt, A. Dietl, N. Kip, U. Mersdorf, T. R. M. Barends, M. S. M. Jetten, J. T. Keltjens and B. Kartal Characterization of anammox hydrazine dehydrogenase, a key N₂-producing enzyme in the global nitrogen cycle, *J. Biol. Chem.*, 2016, **291**, 17077-17092.
26. J. D. Clarke, *Ignition! An Informal History of Liquid Rocket Propellants*, Rutgers University Press, New Brunswick, NJ, 1972.
27. G. LeGoff and J. Ouazzani Natural hydrazine-containing compounds: biosynthesis, isolation, biological activities and synthesis, *Bioorg. Med. Chem.*, 2014, **22**, 6529-6544.
28. C. G. Mowat and S. K. Chapman Multi-heme cytochromes-new structures, new chemistry, *Dalton Trans.*, 2005, , 3381-3389.
29. C. J. White and N. Lehnert Is there a pathway for N₂O production from hydroxylamine oxidoreductase in ammonia-oxidizing bacteria?, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 14474-14476.
30. S. W. Watson and R. G. E. Murray Structure of *Nitrosocystis oceanus* and Comparison with *Nitrosomonas* and *Nitrobacter*, *Am. Soc. Microbiol.*, 1965, , 1594-1609.

Nitrate Ion Reductive Metabolism: Bifurcation Between Five-electron or Eight-electron Reductions

4.1 Two Modes of Biological Reduction of Nitrate Ions

Nitrate ions ($N = +5$ oxidation state) represent the endpoint of the most oxidized form of nitrogen atoms in biology. They are the most abundant forms of inorganic nitrogen in the global oceans. Organisms can utilize nitrate ions in two distinct reductive modes. 1 One is denitrification in which nitrate ions are enzymatically reduced back to N_2 , completing the global cycle of nitrogen 2 (nitrogenase, nitrification, denitrification, Figure 4.1). The stoichiometry of denitrification requires two NO_3^- to be reduced for each end product molecule of N_2 ($N = 0$ oxidation state) to be generated, along with six water molecules. Each nitrogen atom in the two nitrate ions undergoes a net five-electron oxidation. Together, denitrification of two nitrate anions to one molecule of N_2 represent a net ten-electron reduction sequence.

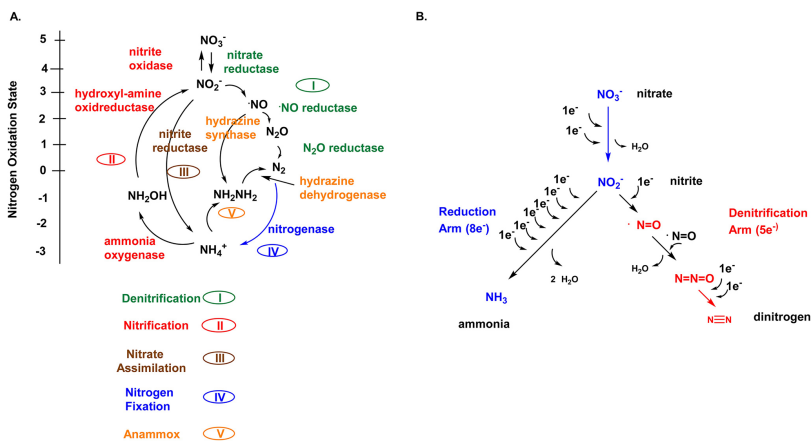


Figure 4.1 (A) Transits between dinitrogen (N_2), ammonia (NH_3) and nitrate ions (NO_3^-) in the arms of the nitrogen biogeochemical cycle. (B) Bifurcation of

The second biological reductive mode undergone by NO_3^- is the net reverse of the ammonia oxidation pathway above with net eight-electron reduction of nitrate to ammonia. Those reductive conversions can occur in a dissimilatory or assimilatory mode. The *dissimilatory* mode uses the reactions involved in conversion of nitrate ion ($\text{N} = +5$) to NH_3 ($\text{N} = -3$) as terminal acceptors for eight electrons removed during central metabolism of organic metabolites [3](#) (e.g. in the citrate cycle). Other bacteria can perform the same series of reductions to ammonia without using the pathway as a major sink for metabolite electrons.

Dissimilatory nitrate reduction is practiced by certain anaerobic bacteria using nitrate instead of O_2 as the terminal acceptor. This logic parallels that of the anaerobic sulfate-reducing bacteria that similarly take sulfate ($\text{S} = +6$) to H_2S ($\text{S} = -2$) as a terminal electron acceptor. This mode of nitrate metabolism is sometimes termed “*nitrate respiration*” to emphasize the substitution of nitrate for O_2 at the end of membrane electron transport chains. [4](#)

In contrast, plants use nitrate ion reduction in *assimilatory* mode. [5](#) It is not the funneling of electrons flowing into nitrate that is the purpose of the pathway. They use O_2 for that purpose, with better thermodynamics. Instead, assimilatory nitrate ion reduction is performed by plants to get the NH_3 end product for conversion to all the organic amino acids, nucleotides and phospholipids that are essential components of life.

The strategy and enzymology of dissimilatory and assimilatory nitrate reductions are equivalent. It is the purposes to which the either the eight electrons removed or the end product NH_3 is used. (We note that the anaerobic bacteria engaging in dissimilatory nitrate reduction for energy purposes do also utilize the NH_3 end product for organifications, discussed in Chapter 6.) It has been estimated that assimilatory nitrate reduction by bacteria algae, fungi and plants accounts globally for annual microbial reduction of 10^{10} tons of nitrate to ammonia for fixation into organic scaffolds. [1](#)

We first take up the conversion of nitrate to ammonia: each nitrate undergoes eight-electron reduction by nitrate dissimilatory and assimilatory reduction where the end product is NH_3 . Second, we turn to denitrification, which results in the end product N_2 . In denitrification each nitrate ion is reduced by five electrons, indicating distinct strategies and mechanisms for the two different redox processes.

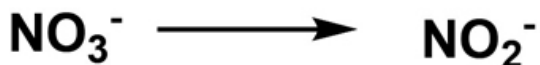
4.2 Nitrate Reduction: NO_3^- to NH_3 for Assimilation: Eight-electron Reduction

The eight-electron reduction pathway of nitrate to ammonia uses distinct enzymes from the ammonia oxidation pathway described in Chapter 3. In particular there are only two enzyme catalysts, the first carrying out a two-electron reduction of nitrate to nitrite. [6](#) The second enzyme funnels six electrons into nitrite without release of any partially reduced nitrogen species before releasing ammonia ([Figure 4.2](#)).

Nitrate to Ammonia

Two enzymes

1. Nitrate reductase. ($2e^-$)



2. Nitrite reductase ($6e^-$)

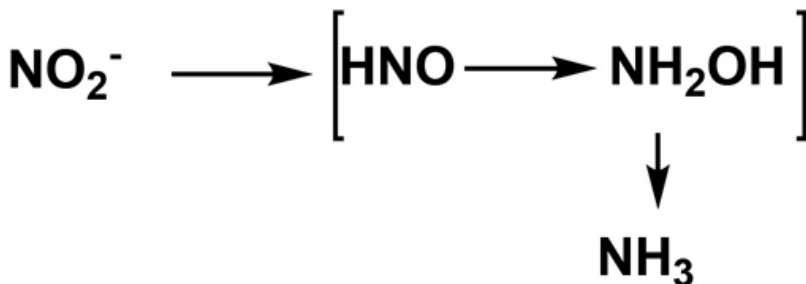
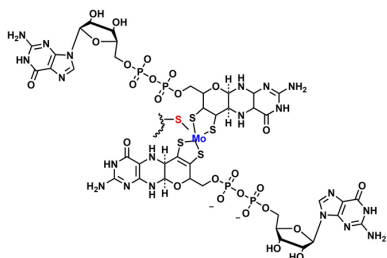


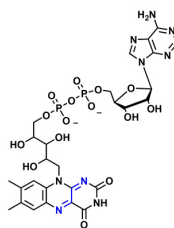
Figure 4.2 The eight-electron reduction of nitrate ion to ammonia is carried out by a two-electron nitrate reductase and a six-electron nitrite reductase.

4.2.1 Nitrate Reductases

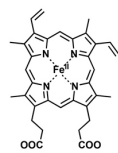
The eukaryotic (plant) and the distinct versions of prokaryotic nitrate reductases have at their structural and catalytic core a molybdopterin cofactor either monomeric, or a dimer of a molybdopterin-guanine dinucleotide (MDG; [Figure 4.3](#)). The extended conformation of the MDG cofactor is shown in the *Escherichia coli* nitrate reductase active site. This is the same molybdopterin guanine dinucleotide cofactor noted in the nitrite oxidase enzyme in the ammonia oxidation pathway, as depicted in Chapter 3 [7](#) ([Figure 3.9](#)). Clearly, the cofactor can catalyze the two-electron transfer in either direction, depending on the apoprotein microenvironment.



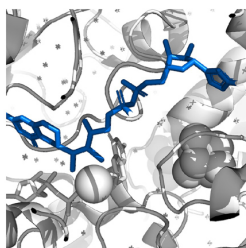
Molybdopterin guanine dinucleotide
(5 thiol ligands around Mo^{VI} Atom)



Flavin Adenine Dinucleotide
(FAD)



Fe^{II} oxidation
state of heme b



E. coli nitrate reductase
with bound MDG cofactor

from pdb 1Q16

Figure 4.3 Three cofactors in nitrate reductase: molybdopterin guanine dinucleotide; flavin adenine dinucleotide, heme-b. The extended conformation of the MDG cofactor in the *Escherichia coli* reductase site is taken from PDB file 1Q16.

In plants, fungi and algae, the nitrate reductase at the start of the assimilatory pathway uses the readily available reduced nicotinamide adenine dinucleotide (NADH) as the external two-electron donor. ⁸ As in many other NADH-utilizing oxidoreductases there is a tightly bound flavin adenine dinucleotide (FAD) coenzyme and also a tetracyclic pyrrole from the cytochrome B family for a total of three bound cofactors and one external NADH substrate (Figure 4.3).

The FAD and the heme B redox cofactors are presumed to feed electrons into the molybdopterin cofactor. ⁶ The route of electron flow starts with a hydride from NADH (obligate two-electron transfer coenzyme form of vitamin B₃) to produce FADH₂. Then, the reduced flavin coenzyme reveals its redox versatility as a two-electron–one-electron redox switch molecule, undergoing two one-electron transfers to get back to the starting FAD state. The first transfer generates the odd electron flavin semiquinone radical (FADH[•]), the second electron out takes the FADH[•] radical back to FAD (Figure 4.4).

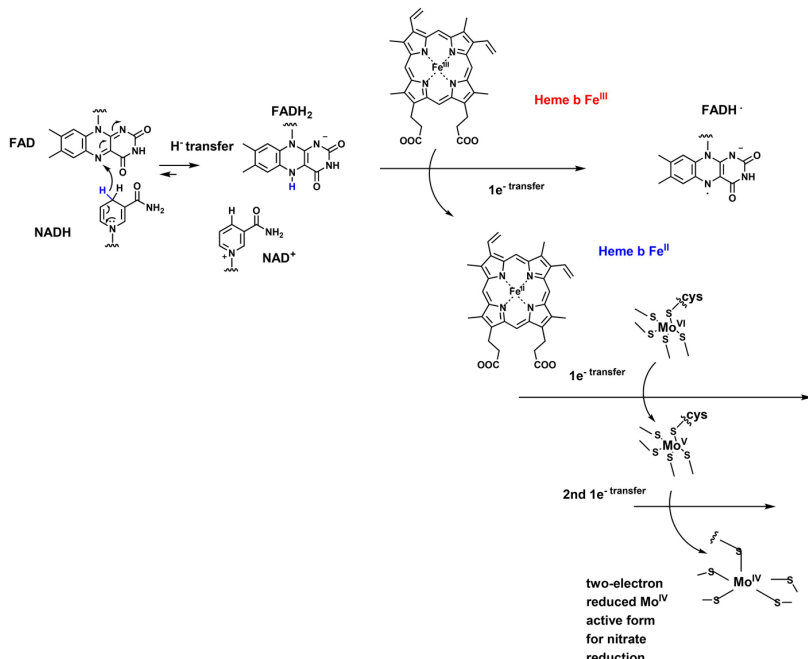
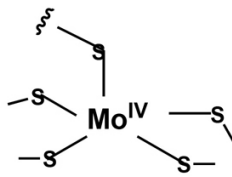


Figure 4.4 Path of electrons from NADH to FAD to heme-b to molybdopterin to substrate nitrate in nitrate reductase.

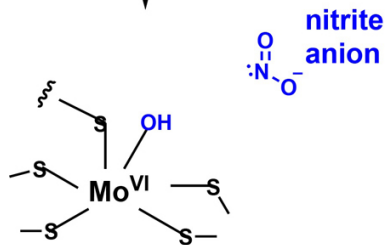
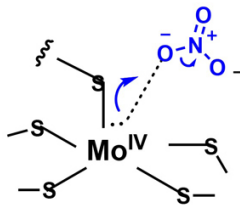
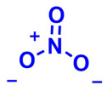
Each single electron, from FADH₂ and then from FADH[•], is delivered to the ferric heme iron atom in the embedded cytochrome b, converting the equatorial Fe^{III} in the tetrapyrrole macrocycle plane to Fe^{II}. Finally, one at a time, the input of two electrons is transferred from the heme iron to the Mo atom in the molybdopterin coenzyme, reducing it first to the Mo^V state and then to the Mo^{IV} state (Figure 4.4). At this point the enzyme–molybdopterin is ready to transfer the two electrons from the reduced molybdenum atom to the substrate nitrate ion. ⁷

In a reaction sequence that is effectively the *reverse* direction of the bacterial nitrite oxidase in the final step of nitrification (Chapter 3), one of the nitrate oxygens forms a nitrate ester to the electrophilic oxygen ligand at the Mo^{IV}. Decomposition in the forward direction involves use of the electron pair in the Mo^{IV} atom for reductive cleavage of one of the three N–O bonds in the tethered nitrate. This step reoxidizes Mo^{IV} back to starting Mo^{VI} oxidation state and produces a nitrite ion and a molecule of water ⁷ (Figure 4.5).



**Mo^{IV} state of
nitrate reductase**

nitrate
anion



**resting Mo^{VI}
oxidation state**

Figure 4.5 Reduced Mo^{IV} is reoxidized to Mo^{VI} as nitrate is reduced to nitrite.

Unlike the above eukaryotes with one form of nitrate reductase for ammonia assimilation, some anaerobic bacteria contain up to three forms of nitrate reductases:

an assimilatory nitrate reductase, a membrane-bound respiratory nitrate reductase (where electrons are funneled from central metabolism *via* NADH to nitrate as a terminal electron acceptor) and a periplasmic nitrate reductase. ⁹ The bacterial nitrate reductases can likewise use NADH (e.g. in *Bacillus subtilis*) as an outside two-electron donor or, in its place, the reduced form of the protein ferredoxin. While FAD and the molybdopterin dimers are requisite cofactors, the b-type cytochrome is absent in the bacterial enzyme, replaced by an Fe-S cluster as an alternate one-electron iron-based electron conduit. The *E. coli* respiratory nitrate reductase is induced during anaerobic growth as a three-subunit enzyme with a heme b type in the NarI subunit, Fe₄S₄ in the NarH subunit and the molybdopterin in the NarG subunit. The electron donor is reduced coenzyme Q (CoQH₂) in the membrane, feeding electrons through NarI to NarH to NarG for nitrate reduction. Clearly, there have been a number of evolutionary efforts to funnel substrate-derived electrons into the molybdopterin *via* distinct one-electron conduits.

The overarching strategy in all organisms that reduce nitrate to nitrite is to deliver two electrons to the Mo^{VI} resting form of the molybdopterin cofactor at the active site to generate the Mo^{IV} species. The Mo^{IV} ligates incoming NO₃, transiently generates a Mo^{IV}-nitrate ester adduct, and then decomposes in the forward direction with two-electron transfer from Mo to the nitrogen of substrate nitrate ion.

4.2.2 Nitrite Reductases

The transfer of the remaining six electrons needed for reduction of a nitrite ion all the way to NH₃ is performed by a single six-electron transfer enzyme, nitrite reductase. ¹⁰ Multiple strategies have been evolved by microbes vs. plants for design and construction of nitrite reductases, depending on whether a one-electron reduction to release nitric oxide in denitrification (Section 4.3 below) or a six-electron reduction to get to ammonia in assimilatory reduction is the physiological goal (Figure 4.1B). The enzymes that catalyze six-electron reduction of nitrate to ammonia do so without any release of intermediate nitrogen oxides along the way.

4.2.2.1 Six-electron Reduction by Plant Nitrite Reductases

We take up the assimilatory spinach nitrite reductase first. In one sense there is analogy to the other six-electron reduction catalysts in nitrogen biogeochemical cycles, the nitrogenases. We emphasize in the overview Chapter 5 that nitrogenase deals only with nitrogen hydride chemistry while nitrite reductase deals with nitrogen oxide and substitutes N-H for N-O bonds during a catalytic cycle. A closer parallel to the plant nitrite reductase strategy and catalytic inventory is provided by bacterial and plant sulfite reductases.

Not only do sulfite reductases carry out comparable six-electron reduction of sulfite ions to H₂S (Figure 4.6), they employ the identical kind of catalytic inventory. Two redox elements in common are: (1) a cubane type of Fe₄S₄ iron-sulfur cluster as a one-electron conductor from the external reductant to (2) a central *siroheme* that is the catalytic center for all six electrons into nitrite. Siroheme (Figure 4.7A) is a variant of a tetrapyrrolic porphyrin macrocycle known as an isobacteriochlorin. (Siroheme is a highly anionic octacarboxy tetrapyrrole macrocycle. Siroheme is purple whereas heme-b is reddish brown.) We discuss the logic and mechanisms of heme and siroheme assembly in Chapter 11.

Nitrite ion

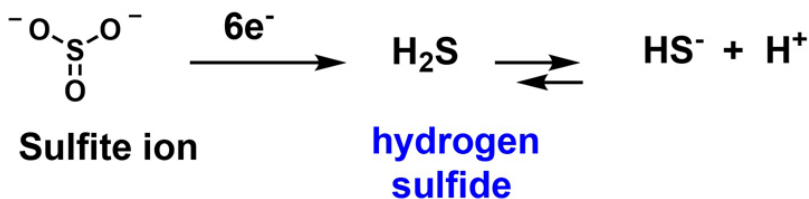


Figure 4.6 Cognate enzyme machinery and redox logic between six-electron nitrite ion and sulfite ion reductases.

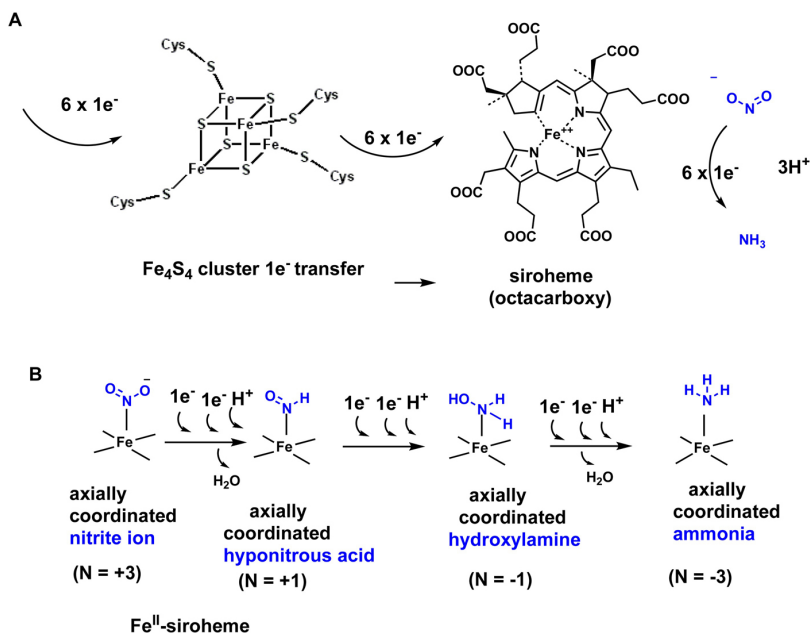


Figure 4.7(A). The two types of iron cofactors in nitrite ion reductase are a feeder Fe₄S₄ cluster and the octacarboxy heme siroheme. (B). Proposal for six single-electron transfers to axially coordinated nitrite-siroheme complex to generate bound intermediates hyponitrous acid and hydroxylamine on the way to NH₃.

The only two enzymes known to contain and use tightly bound siroheme are sulfite reductase and nitrite reductase. The external reductant is ultimately NADH in both enzymes, converted to reduced ferredoxin proteins by NADH ferredoxin oxidoreductases. Electron flow from NADH or ferredoxin to FAD (2e⁻ or 1e⁻ step down) to Fe₄S₄ sets up a conduit to the siroheme. That Fe₄S₄ cluster is connected electronically to the siroheme (Figure 4.7A), allowing efficient flux of electrons one at

a time into the siroheme.

The substrate NO_2^- probably starts every catalytic cycle as the top axial ligand to ferric siroheme. Subsequent one-electron reduction would form the $\text{NO}_2^- \text{Fe}^{\text{II}}$ siroheme as a nitrosyl radical complex. Injection of a second electron would yield a coordinated hyponitrous acid- Fe^{II} complex as depicted in Figure 4.7B. Subsequent consecutive third and fourth electron injections from the external reductant could give an axially coordinated $\text{NH}_2\text{OH-Fe}^{\text{II}}$ -siroheme. Finally, input of the fifth and sixth electron would lead to product NH_3 axially coordinated to the siroheme central Fe atom before its release into solution (Figure 4.7B).

The exact timing of entry of each of the six electrons, the equilibration of electrons between the Fe^{III} and Fe^{II} states of the siroheme, and the electron transfers to axially coordinated nitrogen oxide species are not fully known. 11 The same ambiguities of timing exist for the comparable sulfite reductase intermediates. What is not in doubt is the spectacular efficiency of this pair of six-electron reductases as they respectively convert electrophilic species of nitrite ($\text{N} = +3$) or sulfite ($\text{S} = +4$) to the essential nucleophilic forms of ammonia ($\text{N} = -3$) or hydrogen sulfide ($\text{S} = -2$). It is these low-valent nucleophilic nitrogen and sulfur species, *and only these*, that have the chemical reactivity for incorporation into organic scaffolds [as in glutamate or cysteine formation, respectively (Chapter 6)].

In comparison to the spinach and other plant nitrite reductase logic, bacterial reductases recapitulate multiple strategies, perhaps as early experiments in reductive electron transfer to nitrogen oxides and/or catalysts that route nitrite ions to NH_3 (e.g. the dissimilatory pathway to ammonia). 12 Some bacteria use the iron-sulfur and siroheme paradigm as above for the spinach or algal six-electron nitrite reductases. Others, such as *E. coli*, can elaborate and secrete the enzyme NrfA into the periplasm. 13 This enzyme contains up to five type c heme groups. Four of them are thought to be electron storage units (up to four electrons total) while the fifth heme is thought to bind nitrite ion as an axial ligand to the central iron atom. Reduction of bound nitrite all the way to NH_3 is again accomplished without accumulation or release of any detectable intermediates.

4.3 Denitrification: two NO_3^- to N_2 , Net Ten-electron Reduction

A second major pathway of bacterial nitrate reduction diverts nitrate ions to inert dinitrogen rather than the utilizable ammonia. 2 The denitrifying pathway, practiced by over 50 genera of diverse bacteria (e.g. *Pseudomonas stutzeri*) is found in 10–15% of recoverable bacteria from marine and terrestrial niches. *Denitrification* returns N_2 to the atmosphere as a major arm of the nitrogen biogeochemical cycle in the order of 10^{10} tons of N_2 per year. The enzymatic pathway and diversion of flux from ammonia to N_2 reveals distinct chemical and metabolic logic.

Figure 4.1B indicates that denitrification involves three intermediates on the way to N_2 .

They are, in order of appearance, nitrite ion (NO_2^-) ($\text{N} = +3$ oxidation state), nitric oxide as a gaseous nitrogen radical (N oxidation state = +2), and nitrous oxide (N_2O) ($\text{N} = +1$ oxidation state). The gaseous, inert end product N_2 is defined as the $\text{N} = 0$ oxidation state. Of immediate notice is the presence of nitric oxide ($\text{N} = 0$), the relatively stable nitrogen radical, along with the observation that it is only one electron down from the nitrite ion. Both of the metabolic intermediates nitric oxide and nitrous oxide are gases under physiological conditions (as is N_2) and some portion of them can diffuse out of solution into the air before further reduction. These are

significant pathways to their emission and accumulation as atmospheric greenhouse gases.

4.3.1 Denitrifying Nitrate Reductase

The nitrate reductases in denitrifying bacteria are the same catalysts that initiate the assimilatory and dissimilatory reductions on the way to ammonia noted in [Figures 4.4–4.6](#).¹⁴ These are the same flavin–iron–molybdopterin catalysts used in the first step of the nitrate to ammonia pathway noted above. Indeed, both pathways share the protein architecture, cofactor chemistry and underlying two-electron reduction logic. Only at the next step of processing the nitrite ion does the denitrification pathway diverge from the eight-electron pathway to ammonia.

4.3.2 Denitrifying Nitrite Reductases: One-electron Reductions

Two types of nitrite reductases are known in denitrifying bacteria.¹² Each has very different architectures from the six-electron reduction catalysts in the assimilatory reduction pathway noted above. The denitrifying enzymes carry out only a one-electron reduction of nitrite to nitric oxide ($\text{N}=\text{O}$).

Most prominent are nitrite reductases of the cytochrome cd1 type,¹⁵ found in species such as *Pseudomonas denitrificans* and *Pseudomonas aeruginosa*. These are dimeric enzymes, with one type c cytochrome in one subunit and one modified, d-type cytochrome in the other subunit ([Figure 4.8A](#)). The heme d modification is a conversion of two of the four pyrrole rings in the tetrapyrrole macrocycle to pyrrolidone rings with acetic acid side chains, altering electronics and macrocycle geometry.^{16,17} This heme d central iron atom is the site for axial ligation of the nitrite anion.

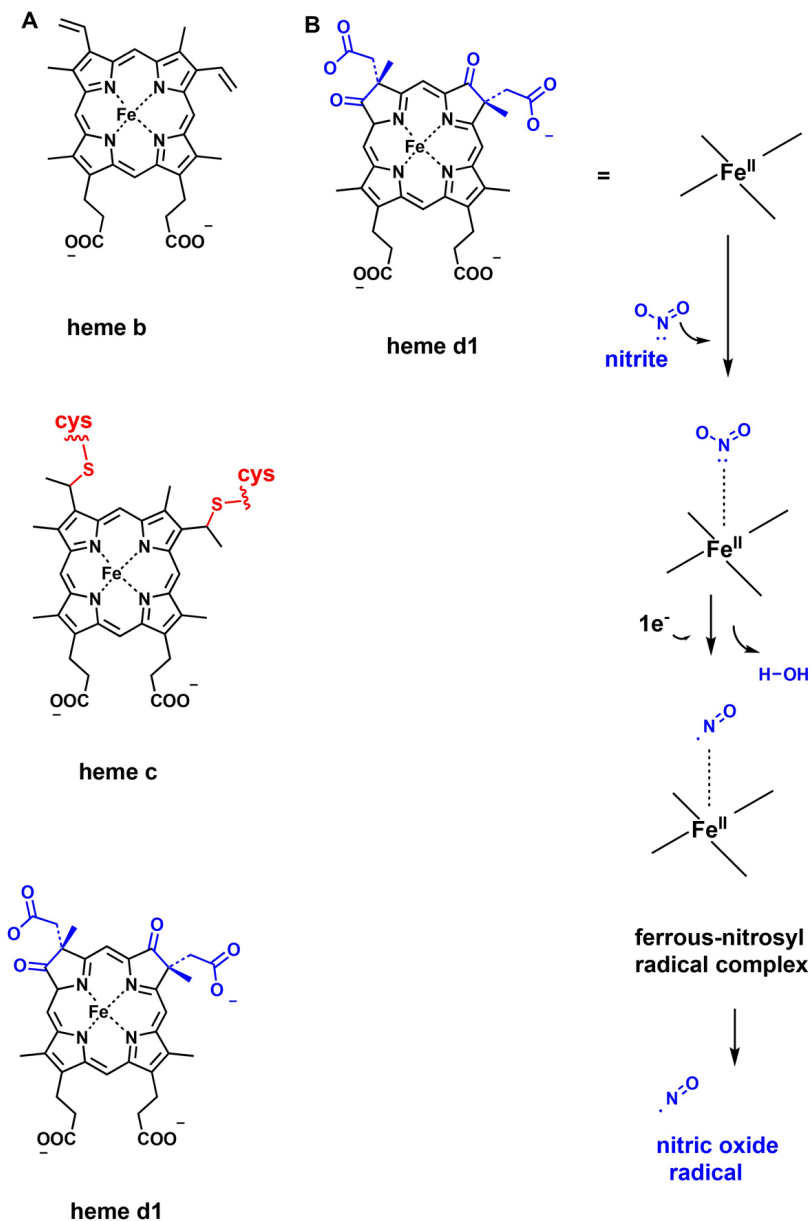


Figure 4.8(A) Three heme variants: heme-b, bis thioether-heme-c, bis-pyrrolidone-heme d1. (B) One-electron reduction of nitrite ion to nitric oxide radical while axially coordinated to heme-d1.

Net one-electron reduction to the axially coordinated nitric oxide radical $\text{N}=\text{O}^\bullet$ is then achieved. It may be that a ferric ion (Fe^{III}) heme d is first reduced to Fe^{II} , then a nitrite ion binds axially. Input of a second electron could produce the one-electron reduced $\text{Fe}^{\text{II}}\text{-NO}$ product complex (Figure 4.8B). In any case, this is a *net one-electron reduction* of a modified heme-coordinated nitrite in sharp contrast to the six-electron reduction of a modified heme-nitrite seen in the full reduction to ammonia.

The second and less frequently observed variant of denitrifying nitrite reductases are copper redox enzymes (Figure 4.9A) found in such bacteria as *Neisseria gonorrhoeae* and *Neisseria meningitidis*. A paradigm of this strategy is found in the trimeric NirK enzyme, with a pair of copper ions, a type I and a type II-copper center, in each subunit. 18 The type II copper center is the binding site for a nitrite ion (Figure 4.9B). An external copper protein, such as azurin, serves as an external single-electron donor to the copper type I center. That electron then flows to the type II copper center, reducing bound nitrite ion by one electron to give [NO].

**Nirk type copper-containing
nitrite to nitric oxide reductase**

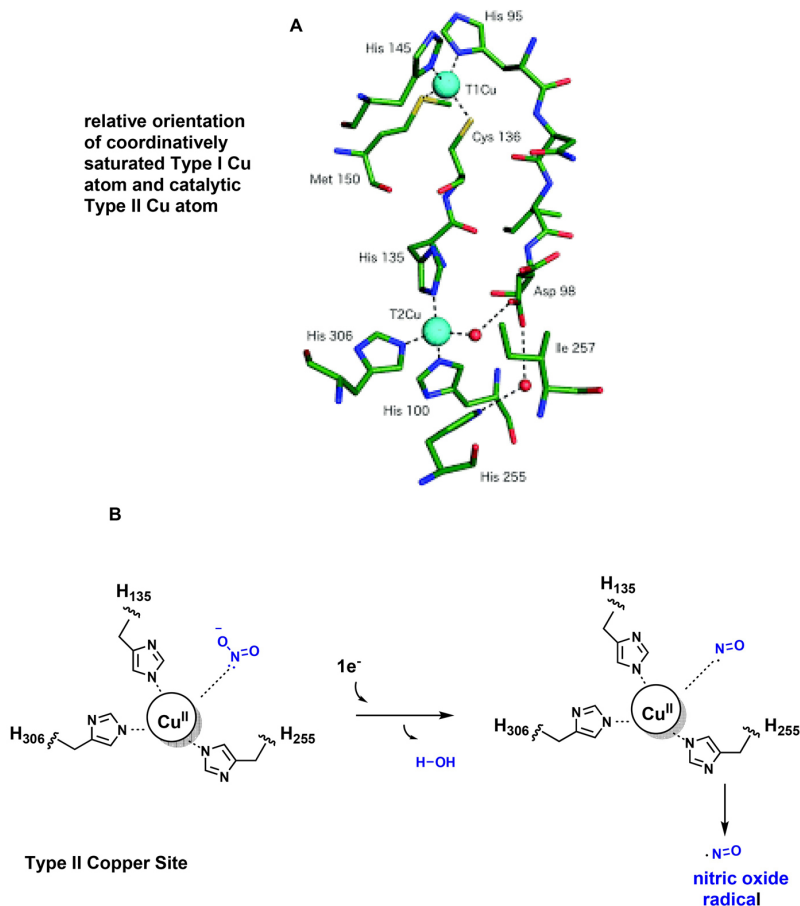


Figure 4.9 Alternative catalytic logic and machinery in a copper-dependent bacterial nitrite to nitric oxide reductase. (A) Orientation of type 1 (T1) structural and catalytic copper centers [type 2 (T2)]; Reproduced from ref. 18 with permission from the Royal Society of Chemistry. (B) Proposed ligation of nitrite and one electron reduction to nitric oxide radical at type 2 copper site.

The nitric oxide radical shows up as a diffusible or heme-bound metabolite in three of the four inorganic nitrogen redox pathways: in ammonia oxidation to nitrate, in the anammox pathway to N_2 and now in denitrification. Such nitric oxide metabolic ubiquity indicates that it may have been a significant ancient form of nitrogen oxide as

the atmosphere became aerobic.

4.3.3 Nitric Oxide Reductase in Denitrifying Bacteria

The third of the four enzymes in the bacterial denitrification pathway is nitric oxide reductase, converting the nitric oxide radical into the product nitrous oxide 19 (also known as laughing gas from its historical use as a dental anesthetic agent). During the denitrification pathway both gaseous nitric oxide and nitrous oxide can escape into the atmosphere as lost, partially reduced intermediates. Each is a problematic greenhouse gas, adding to global warming.

Note from the stoichiometry in Figure 4.10 that this enzyme reductively dimerizes two molecules of nitric oxide radicals to N_2O product. The nitrogen redox state in $N=O$ is $N = +2$ while in nitrous oxide ($N_2=O$) $N = +1$. A single electron has been added to each of the two nitric oxides as $N_2=O$ is formed. During the coupling of two nitric oxide radicals a new N–N bond is formed and an N–O bond is broken. The possible timing is addressed below.

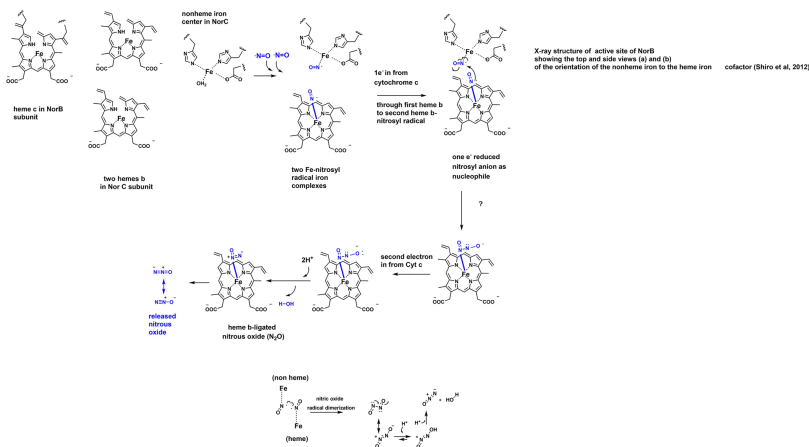


Figure 4.10 The NorC nitric oxide to nitrous oxide reductase has two heme-b-type and one mononuclear nonheme iron site. One heme acts as an electron conduit, the second heme is the binding site of the axial ligand of one nitric oxide while the second nitric oxide ligates to the nonheme iron. The N–N bond forming polarity is suggested to involve attack of the heme-coordinated nitroxyl radical on the nonheme-coordinated nitric oxide radical: homolytic N–N bond formation. (B). Active site orientation of the mononuclear nonheme iron center to the heme iron center in NorC. Reproduced from ref. 23 with permission from The Royal Society, Copyright 2012.

In the denitrifying bacterium *Pseudomonas aeruginosa*, the enzyme is a two-protein complex of a periplasmic NorC subunit that contains a c type cytochrome, functioning as the proximal one-electron donor to the NorB subunit. NorB contains two heme b-type tetrapyrrole macrocycles (Figure 4.8). One heme b acts as electron conduit from NorC to the second heme that is part of a binuclear iron center, comprising this second b-type heme and a mononuclear nonheme iron known as Fe_b (Figure 4.10A). Among the many iron-based enzymes in the four pathways of inorganic nitrogen biogeochemical cycling, this is the only heme–mononuclear nonheme iron redox couple. 20 One electron at a time flows from an external reductant, such as the protein azurin, to the NorC heme, into the first of the b-type hemes in NorB and finally into

events to produce ligated N_2 at the tetranuclear copper.

The tetranuclear copper site, sometimes termed the Cu Z site in the reference literature, is novel structurally and is the locus of N_2O binding and reduction. As depicted in Figure 4.11A (Figure 3 of ref. 22), one can find a Cu_4S_2 array in the *Pseudomonas stutzeri* reductase but a Cu_4S_1 array in the *P. denitrificans* form of the reductase. The thought is that the Cu_4S_1 form is catalytically active. N_2O binding has been modeled onto the Cu_1 and Cu_4 edge of the Cu Z tetranuclear copper cluster and a possible reaction cycle is shown in Figure 4.11B. 22 The N_2O molecule bound between the two copper ions can undergo one-electron reduction with N–O bond cleavage to release gaseous N_2 . Then, a second proposed one-electron input per reduction leads to release of H_2O , with turnover rates approaching 8000 catalytic events per minute.

4.4 Overview of the Metallocenter Inventory in the Denitrification Pathway

For each pair of nitrate ions shunted down the four-enzyme pathway to N_2 , the bacteria that carry out global denitrification employ three redox active metals, molybdenum, iron and copper to mediate the net five-electron reduction (Figure 4.12). They also utilize one organic cofactor, the riboflavin-based FAD. The iron and copper systems are one-electron conduits, the FAD is a two-electron–one-electron step down switch. The molybdenum also functions as a two-electron–one-electron redox switch.

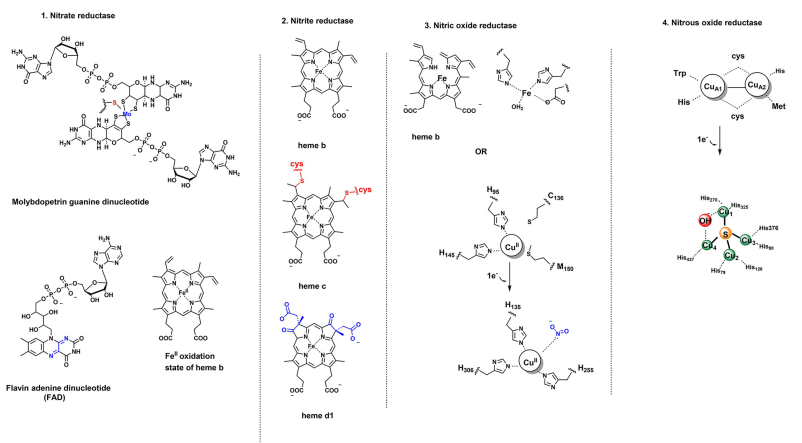


Figure 4.12 Diverse iron- and copper-based redox cofactors in the enzymes involved in the two distinct pathways for nitrate ion reduction: assimilatory eight-electron reduction to ammonia or ten-electron reduction to dinitrogen (N_2).

From the point of view of producing N_2 , two nitrate ions have to be reduced and combined to form the strong N–N triple bond, so the denitrification process removes a total of ten electrons from the two nitrate ions to yield one N_2 .

In the first step it is the directed cycling of Mo^{VI} to Mo^{IV} that creates the reactive center for Mo–nitrate ester formation. Then, elimination of water with two-electron input to the liganded nitrate, occurs with the driving force provided by simultaneous reoxidation of Mo^{IV} to the starting Mo^{VI} oxidation state. In contrast to the FeMo cofactor of nitrogenase where the Mo^{VI} atom does not participate directly in the redox

process, the Mo in the molybdopterin cofactor of nitrate reductase is the primary redox center.

The first three enzymes – nitrite reductase, nitric oxide reductase, and nitrous oxide reductase – all employ one or more heme b type tetrapyrrole macrocycles, reflecting the ubiquitous adaptation of the heme macrocycle properties for intraprotein and interprotein single-electron transfers. In nitrate reductase the heme b interfaces between reduced FADH₂ and the molybdenum atom in molybdopterin. In nitrite ion reductase, three heme variants are employed and again heme b is an intermediate conductor of one electron (Fe^{III} to Fe^{II} and back to Fe^{III}), this time between exogenous cytochrome c donor and the heme d1 that is the site of nitrite single-electron reduction to the nitric oxide radical (Figure 4.12). In the third enzyme the heme b is now directly part of the catalytic action via axial nitrosyl radical complex, interfacing with a nonheme mononuclear iron center also as a nitroxyl radical complex for N-N bound coupling. (A fourth heme variant is the P460 heme found in the ammonia oxidation and anammox pathways described in Chapter 3).

Copper is utilized in two of the four enzymes in some bacterial denitrification pathways. The nitric oxide reductase appears to have evolved twice. One route has yielded the heme b–nonheme mononuclear iron redox pair at the catalytic center. The other evolutionary path has led to a two-copper enzyme where each copper functions as a mononuclear copper. Of the two coppers, one is coordinatively saturated and acts as a one-electron relay from an exogenous reductant. The second copper center binds nitrite and is the proximal one-electron reductant of nitrite ions.

The second appearance of copper is in the fourth and final enzyme of the denitrification pathway. Remarkably, each catalytic subunit contains six copper atoms, organized as a binuclear copper center and a tetranuclear copper center (Figure 4.12). The binuclear center is the intermediate electron conduit while the tetranuclear copper array, anchored by a central sulfide ion, is the catalytic center. This Cu₄S₁ center is unique structurally, while functionally it mediates input of two single electrons to N₂O coordinated to one of the four coppers. The result is controlled release of a product water molecule by cleavage of one of the two N=O bonds of N₂O with conversion of the N=N bond in N₂O to the strong N-N triple bond in N₂.

In sum, one can marvel at the diverse array of inorganic (and organic redox platforms) that unicellular denitrifying bacteria bring to bear on the net ten-electron reduction of two NO₃⁻ to one molecule of dinitrogen (N₂). This nitrogen redox chemistry surpasses any of the redox transformations at nitrogen in the organic nitrogen scaffolds used throughout nucleic acid, protein and intermediary metabolism.

References

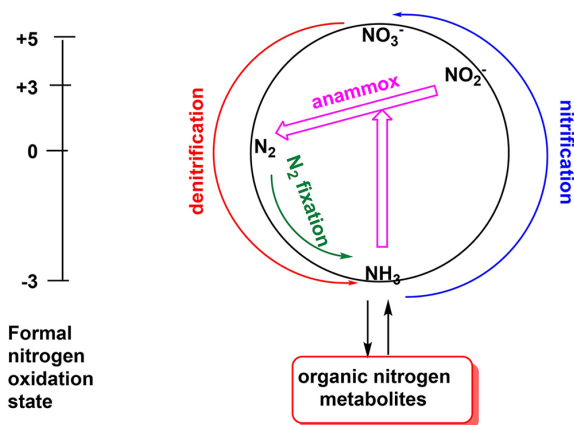
1. C. Moreno-Vivian, P. Cabello, M. Martinez-Luque, R. Biasco and F. Castillo Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases, *J. Bacteriol.*, 1999, **181**, 6573-6584.
2. W. G. Zumft Cell biology and molecular basis of denitrification, *Microbiol. Mol. Biol. Rev.*, 1997, **61**, 533-616.
3. S. An and W. Gardner Dissimilatory reduction to ammonium (DNRA) as a nitrogen link versus denitrification as a sink in a shallow estuary (Laguna Madre/Baffin Bay, Texas), *Mar. Ecol. Prog. Ser.*, 2002, **237**, 41-50.
4. J. Simon and M. Klotz Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations, *Biochim. Biophys. Acta*, 2013, **1827**, 114-135.
5. G. Xu, X. Fan and A. J. Miller Plant nitrogen assimilation and use efficiency, *Annu. Rev. Plant Biol.*, 2012, **63**, 153-182.

6. W. H. Campbell Nitrate reductase structure, function and regulation: bridging the gap between biochemistry and physiology, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1999, **50**, 277-303.
7. G. Schwarz, R. Mendel and M. W. Ribbe Molybdenum cofactors, enzymes, pathways, *Nat. Rev. Microbiol.*, 2009, **460**, 839-847.
8. K. Fischer, G. Barbier, H.-J. Hecht, R. R. Mendel, W. H. Campbell and G. Schwarz Structural basis of eukaryotic nitrate reductase: crystal structures of the nitrate reductase active site, *Plant Cell*, 2005, **17**, 1167-1179.
9. C. Moreno-Vivian, P. Cabell, M. Martinex-Luque, R. Blasco and F. Castillo Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases, *J. Bacteriol.*, 1999, **181**, 6573-6584.
10. L. B. Maia and J. Moura How biology handles nitrite, *Chem. Rev.*, 2014, **114**, 5273-5357.
11. O. Einsle, A. Messerschmidt, R. Huber, P. M. H. Kroneck and F. Neese Mechanism of the six electron reduction of nitrite to ammonia by cytochrome c nitrite reductase, *J. Am. Chem. Soc.*, 2002, **124**, 11737-11745.
12. C. Sparacino-Watkins, J. F. Stolz and P. Basu Nitrate and periplasmic nitrite reductases, *Chem. Soc. Rev.*, 2014, **43**, 676-706.
13. H. Wang and R. P. Gunsalus The *nrfA* and *nrfB* nitrite reductase operons in *Escherichia coli* are expressed differently in response to nitrate than nitrite, *J. Bacteriol.*, 200, **182**, 5813-5822.
14. C. Coelho and M. Romão Structure and mechanistic insights on nitrate reductases, *Protein Sci.*, 2015, **24**, 1901-1911.
15. C. D. Richter, J. W. Allen, C. W. Higham, A. Koppenhofer, R. S. Zajicek, N. S. Watmough and S. J. Ferguson Cytochrome cd1, reductive activation and kinetic analysis of a multifunctional respiratory enzyme, *J. Biol. Chem.*, 2002, **277**, 3093-3100.
16. S. C. Baker, N. F. Saunders, A. C. Willis, S. J. Ferguson, J. Hadju and V. Fulop Cytochrome cd1 structure: unusual haem environments in a nitrate reductase and analysis of factors contributing to beta propeller folds, *J. Mol. Biol.*, 1997, **269**, 440-455.
17. S. Bali, A. D. Lawrence, S. A. Lobo, L. M. Saraiva, B. T. Golding, D. J. Palmet, M. J. Howard, S. J. Ferguson and M. J. Warren Molecular hijacking of siroheme for the synthesis of heme and d1 heme, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 18260-18265.
18. S. Horrell, D. Kekilli, R. W. Strange and M. A. Hough Recent structural insights into the function of copper nitrite reductases, *Metallomics*, 2017, **9**, 1470-1482.
19. J. Hendriks, A. Oubrie, J. Castresana, A. Urbani, S. Gemeinhardt and M. Saraste Nitric oxide reductases in bacteria, *Biochim. Biophys. Acta*, 2000, **1459**, 266-273.
20. J. P. Collman, Y. Yang, A. Dey, R. A. Decreau, S. Ghosh, T. Ohta and E. I. Solomon A functional nitric oxide reductase model, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15660-15665.
21. Y. Shiro, H. Sugimoto, T. Tosha, S. Nagano and T. Hino Structural basis for nitrous oxide generation by bacterial nitric oxide reductases, *Philos. Trans. R. Soc., B*, 2012, **367**, 1195-1203.
22. C. Carreira, S. Pauleta and I. Moura The catalytic cycle of nitrous oxide reductase-the enzyme that catalyzes the last step of denitrification, *J. Inorg. Biochem.*, 2017, **177**, 423-434.
23. Y. Shiro, H. Sugimoto, T. Tosha, S. Nagano and T. Hino, *Philos. Trans. R. Soc., B*, 2012, **367**, 1195-1203.

Chemical Strategies Underlying the Biological Flux of Inorganic Nitrogen Species

Inorganic nitrogen metabolism is about redox biology, putting electrons into N_2 and NO_3^- , or taking electrons out of NH_3 . The physiological web for nitrogen intersects with the biogeochemical cycle diagrams ([Figure 5.1A](#)) where the emphasis is on inorganic nitrogen cycling [1](#) with the organic nonvolatile nitrogen bio-universe sequestered in a small unexamined box of nonvolatiles. The nine chapters in the next section of this volume take up that physiological organic nitrogen web. [Figure 5.1B](#) provides an alternate depiction of the nitrogen cycle with the four kinds of inorganic nitrogen-metabolizing soil bacteria described in chapters 2–4 (minus the marine anammox bacteria): nitrogen fixers, ammonifiers, nitrifiers and denitrifiers. The diagram also shows that both ammonia and nitrate ions are taken up by plant roots to be distributed to other plant tissues. The nitrate ions undergo assimilatory reduction back to ammonia for the organification reactions described in Chapter 6.

A Nitrogen cycle with organic nitrogen as entry and exit point for NH_3



B Redox biocycling of N_2 by bacteria and diversion of nitrate for uptake by plant roots

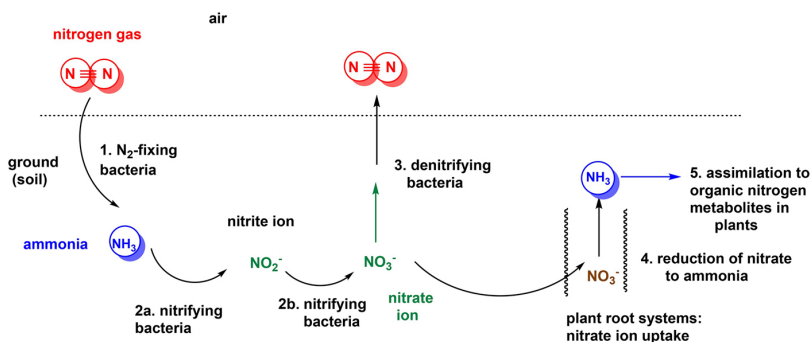


Figure 5.1(A). Nitrogen cycle with NH_3 as the exit point to the universe of organic nitrogen metabolites. (B) The roles of four types of inorganic nitrogen-metabolizing soil bacteria in nitrogen cycling and uptake of both NH_3 and NO_3^- by plant roots.

As we summarize points of underlying logic for microbes and plants that handle the three major inorganic nitrogen compounds, dinitrogen gas, nitrate ions and ammonia, we start from the premise that nitrogenase-mediated reduction of N_2 to NH_3 is the origin of all nitrogen metabolism, both inorganic and organic nitrogen scaffolds. [2:3](#) The anaerobic nitrogen fixing bacteria hold undisputed pride of place, followed by nitrifying bacteria to produce the nitrate that is the main form of inorganic nitrogen taken up by plants and subsequently assimilated reductively back down to NH_3 . Finally, it is NH_3 , with its nucleophilic and basic nitrogen atom, that serves as the cross over metabolite into organic chemical biology. Herbivores live on plants and, secondarily, so do carnivores. Humans are thus two to three steps, depending on diets, away from the elementary fixation of N_2 into metabolically useful forms.

5.1 Traversing the Full Range of Nitrogen Redox States

In their practice of inorganic nitrogen chemical biology, bacteria, fungi and plants employ the full range of nitrogen oxidation states. Figure 5.2A serves as a reminder that inorganic nitrogen biology happens between formal oxidation states of $N = -3$ for ammonia to $N = 0$ for dinitrogen, to $N = +5$ for the nitrate anion. Ammonia (NH_3) contains the most basic and nucleophilic nitrogen atom at the $N = -3$ oxidation state. The $N = 0$ oxidation state of N_2 with its robustly strong N–N triple bond is effectively inert except at the active site of nitrogenase. The nitrogen in the nitrate anion ($N = +5$) has a permanent positive charge and is nonnucleophilic. On the other hand, the nitrite ion (where $N = +3$, two electrons down from the nitrate anion) does have an available lone pair on nitrogen for reactivity.

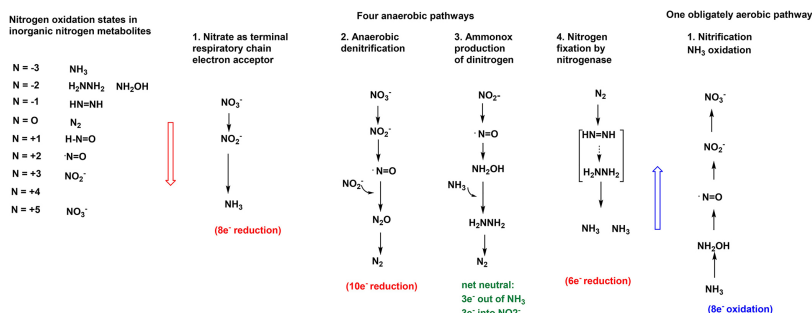


Figure 5.2 Schematic of five redox pathways for inorganic nitrogen metabolites with intermediate N–O and N–H substrates listed. Nitrogen oxidation states are listed. Only one of the five pathways is present in aerobic organisms.

Figure 5.2 also shows inorganic nitrogen metabolites at the corresponding $N = -2$ (hydrazine, hydroxylamine), and the $N = -1$ (diazene states). On the more positive side from N_2 are the inorganic metabolites nitric oxide (N=O where $N = +2$), and nitrite and nitrate anions. The H-N=O hyponitrous acid would occupy the $N = +1$ slot and it is thought to be a metal-bound intermediate in the anammox reaction pathway. There is no corresponding $N = +4$ inorganic nitrogen oxide that is a recognized free metabolic intermediate.

In contrast to the microbes and plants putting inorganic nitrogen through its full eight-electron redox spectrum, vertebrates and mammals interconvert the great bulk of organified nitrogen scaffolds while keeping nitrogen atoms at the $N = -3$ oxidation state. That metabolic regime takes advantage of the basicity and nucleophilicity of amine nitrogens for most of amino acid, urea, protein, RNA and DNA metabolism. *N*-oxygenations do appear in microbial siderophores routes and in plant nitrile biosynthetic pathways as we shall examine in Chapters 12 and 13.

5.2 Three Reductive Metabolic Pathways, One Oxidative Pathway and One Mixed Pathway for Nitrogen Inorganics

Figure 5.2 lays out the five major redox pathways for the above three inorganic nitrogen anchor compounds, discussed in Chapters 2–4. The first four are anaerobic. Molecular oxygen (O_2) is not required and is actually deleterious, causing unwanted competitive oxidation of many of the obligate one-electron transfer coenzymes in these anaerobic pathways. Two of the pathways involve enzymatic reduction of nitrate anion, to distinct fates.

Some anaerobic bacteria use nitrate in place of O_2 as the terminal electron acceptor for the cell membrane electron transport chains (the leftmost pathway in Figure 5.2). As (organic) metabolite electrons fall through the potential drop down to nitrate, the thermodynamics are favorable and allow anisotropic pumping of protons across the cell membranes. Buildup of the proton chemical potential difference across the membranes allows ATP synthesis as the electrons flow back down their chemical gradient. Plants likewise conduct nitrate reduction to ammonia, but for assimilation into organic frameworks. The constituent reductases are presumably in microaerophilic to anaerobic local milieus within plant cells and tissues (e.g. roots) to avoid off pathway oxidation of air-sensitive redox components (see below). These are eight-electron reductions, carried out by only two enzymes, nitrate reductase for the first two electrons in, nitrite reductase for the last six (see below).

The second anaerobic reduction pathway (number 2 in Figure 5.2) is practiced by denitrifying bacteria 4 that convert two molecules of nitrate to N_2 . This is the process that returns hundreds of thousands of tons per year of N_2 back to the global atmosphere. It is the reaction sequence that closes and completes the global biogeochemical cycle of nitrogen flow. Each nitrate ion reduced to one half of the N_2 pathway product undergoes five-electron reduction in a series of three enzymatic steps. To get to a full N_2 molecule from two nitrate anions requires ten-electron input.

The third anaerobic reductive pathway (number 4 in Figure 5.2) is the one that enables organisms to use nitrogen compounds, the fixation of N_2 by net six-electron reduction to two molecules of ammonia, the subject of the second half of Chapter 2. All three of these occur in anaerobes, presumably perfected early in evolution, perhaps before cyanobacterial-mediated oxygenation of the planetary atmosphere was completed.

One pathway in Figure 5.2 (the rightmost one) is different from the first three. This is the nitrification pathway, 5 oxidative rather than reductive, taking eight electrons out of NH_3 on the way to NO_3^- . Electrons flow out from NH_3 into the membrane respiratory chain of those bacteria, rather than flowing into nitrate as in pathway 1. Obligate aerobes run this four-enzyme process (Chapter 3). In two of the four enzymatic steps, O_2 is cosubstrate and donor of an oxygen to the increasingly oxygenated nitrogen framework. We comment on the relative thermodynamics below.

The fifth pathway (number 3 in Figure 5.2) is the most recently elucidated anammox pathway. 6 It joins the first three as occurring in obligately anaerobic bacteria. By coupling NH_4^+ and NO_2^- to N_2 and two H_2O , the anammox pathway avoids both biological formation of nitrate and its subsequent three-step denitrifying conversion to N_2 . Intriguingly, from a redox point of view in comparison to the other four pathways of Figure 5.2, anammox uniquely has *both* an oxidative arm and a reductive arm. The nitrate ion substrate gets reduced by three electrons and the NH_3 substrate gets oxidized by three electrons. Their nitrogen oxidation states meet in the middle as NO_2^- goes from +3 to 0 and NH_3 goes from -3 to 0. The timing and mechanism for getting to two nitrogen atoms in N_2 , both at oxidation state $N = 0$ provides an intriguing look into nitrogen oxide and nitrogen hydride chemical biology and how to mix them.

An even closer look at the three enzymes in the anammox pathway reveals that the nitrite ion actually undergoes a four-electron reduction by the first two enzymes: nitrite to nitric oxide and then to hydroxylamine (a transit of nitrogen oxidation states

from (+3) to (+2) to (−1) at the NH_2OH level from a net one- and then three-electron reduction sequence). From the perspective of odd electron vs. even electron chemistry of nitrogen taken up in Chapter 14, this two-enzyme stretch touts the ability of microbial enzymes to deal with nitrogen oxide radical species. (In complementary fashion, the hydrazine to dinitrogen transform is a four-electron oxidation, see below).

The N–N bond-forming step in the second enzyme, hydrazine synthase or hydroxylamine oxidoreductase 7 takes NH_3 at oxidation state $\text{N} = -3$, and NH_2OH at $\text{N} = -1$ to form hydrazine where both nitrogens are now at the $\text{N} = -2$ state, no net redox change.

Given the four electrons going into the intermediates to get from nitrite to hydroxylamine and then to hydrazine, the third anammox pathway enzyme, hydrazine dehydrogenase, 8 must do a net four-electron oxidation to take both hydrazine nitrogens ($\text{N}, \text{N} = -2, -2$) to dinitrogen ($\text{N}, \text{N} = 0, 0$). This is accomplished in two $2e^-$ steps with ligated diazene ($\text{N}, \text{N} = -1, -1$) as an intermediate.

5.3 Nitrogen Oxide vs. Nitrogen Hydride Chemical Biology

Of the five pathways that constitute the (inorganic) nitrogen biogeochemical cycle in Figure 5.2, three of them involve nitrate, the trioxxygenated, most highly valent form of nitrogen ($\text{N} = +5$). In two of the pathways nitrate is the starting substrate for reduction, while in the nitrification–ammonia oxidation pathway nitrate is the end product. During the reduction processes trioxxygenated nitrate undergoes three net reductive dehydrations, replacing each N–O bond with an N–H bond sequentially. At the end the anaerobic microbes have switched from nitrogen-oxide to nitrogen-hydride chemistry completely.

In the other direction, ammonia oxygenation during nitrification, the switch is progressively in the other direction from nitrogen-hydride to nitrogen-oxide chemistry. We noted in Chapter 3 that the first two of the N–O bonds in nitrification are introduced by N-oxygenases (These are harbingers of Chapters 12 and 13). The third oxygen as nitrite goes to nitrate comes from water. Reductive N–O cleavage by net dehydration is a fundamentally different mechanistic chemical logic from radical-mediated N–O bond formation from high valency oxo-iron enzyme intermediates. The oxygenations are one-electron transfer steps (see Chapter 14), with homolytic (radical) bond formation. In contrast, the N–O bond cleavage during dehydrative reductions is probably heterolytic, ejecting $[\text{OH}^-]$ rather than $[\text{OH}]$.

The anammox pathway is an informative twist on conversion from nitrogen-oxide to nitrogen-hydride chemical biology in being both intermolecular and using two substrates that are six oxidation states apart. The first half of the pathway deals with nitrite ion as substrate and its three-electron reduction to strip away one of the two oxygens from nitrite. The crossover metabolite from nitrogen-oxide to nitrogen hydride chemical biology is hydroxylamine, bearing one N–OH and two N–H bonds. The product from the remarkable enzyme hydroxylamine reductase 7·8 is hydrazine (Figure 5.3). The N–O bond in hydroxylamine has been replaced by the N–H bond in hydrazine in a nifty piece of displacement (dehydration) chemistry using NH_3 as a nitrogen nucleophile. The last piece of the anammox pathway is purely N–H chemistry (hydrazine to diazene to dinitrogen).

inorganic nitrogen metabolite redox chemistry as the planet oxygenated. ⁹ One measure of its lifetime is that nitric oxide can escape from the denitrification pathway in microbes and volatilize into the atmosphere where, *inter alia*, it can behave as a greenhouse gas, leading to global warming.

In contrast to the near ubiquitous role of nitric oxide in inorganic nitrogen chemical biology, nitrous oxide (N₂O) is a metabolic singleton ¹⁰ (Figure 5.2). It is formed from nitric oxide dimerization in the penultimate step of anaerobic bacterial denitrification. It, too, can escape before further reduction and volatilize to become another atmospheric greenhouse gas contaminant. The absence of N₂O from the assimilatory nitrate reduction pathway is a reminder that they are indeed distinct in logic, outcome, scope of electrons traversed and inventory of metallo-cofactors utilized.

5.4 From One to Six Electrons Transferred in a Given Enzyme Site

The great bulk of redox reactions in intermediary metabolism involve net two-electron transfers. The reaction manifolds may be hydride transfers (2e⁻ at a time *via* nicotinamide coenzymes) or two one-electron transfer steps (*e.g.* flavin cofactors) but the net stoichiometry of organic functional group oxidations in metabolism are two-electron processes.

By comparison Figure 5.4 demonstrates robust diversity in the net redox stoichiometries of enzymes in the five pathways of inorganic nitrogen metabolism under evaluation in Chapters 2–4. Enzymes that transfer one, two, three, four or six electrons participate. A one-electron transferase example is the nitrite reductase in the anammox pathway, one of the enzymes that gives the nitric oxide radical noted above. The traditional two-electron stoichiometry is seen in nitrate reductase action in the nitrate assimilatory reductase pathway, noted in the molybdopterin oxidation of the embedded Mo^{IV} to Mo^{VI}.

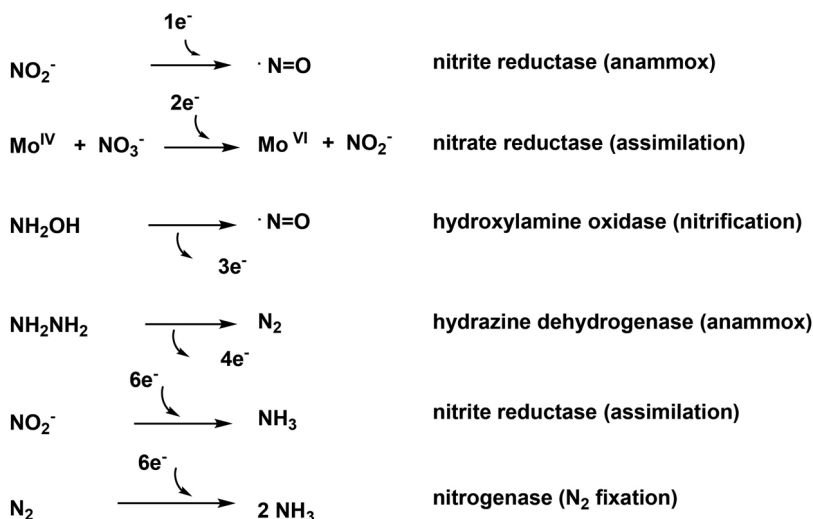


Figure 5.4 The inorganic nitrogen redox pathway enzymes catalyze one-, two-, three-, four- or six-electron transfer processes before release of products.

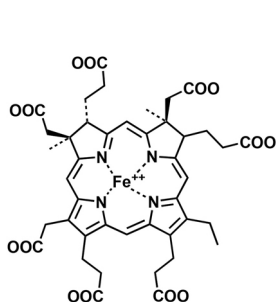
Much less common is the net three-electron oxidation of hydroxylamine. Again, a nitric oxide radical is the product, presumably passing through a two-electron-reduced metal-hyponitrous acid liganded species. We have noted that hydrazine dehydrogenase [8](#) catalyzes the last double step (four-electron oxidation) in the anammox production of N_2 . Finally, two examples of enzymes that sequentially add six electrons to iron cofactor bound NO_2^- or N_2 pump in six electrons before any product is released. We note the distinction in the two iron cofactors in the next section. Four of the six enzymes in [Figure 5.4](#) are reductases, putting electrons into the nitrogen substrate. The other two, hydroxylamine oxidase and hydrazine dehydrogenase, instead take electrons out of the nitrogen substrates. Clearly, the redox cofactor inventory in these nitrogen metabolic pathways enable a dramatically diverse flow of electrons.

5.5 Redox Cofactor Inventories

The redox cofactors are most commonly the two biologically available redox metals, the Fe^{II} and Fe^{III} and the Cu^I and Cu^{II} atoms in different coordination environments. They are typically one-electron transfer reagents.

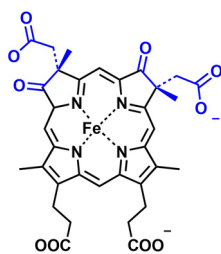
The eleven enzymes of the five inorganic nitrogen metabolism pathways in [Figure 5.2](#) use iron atoms in four distinct architectural and functional contexts, copper ions in three distinct contexts and molybdenum in two cofactor scaffolds. This inventory immediately gives a sense of the span of redox-active metal centers brought to bear as cofactors for specific nitrogen-oxide and nitrogen-hydride types of transformations.

Without going back over specific catalysts evaluated in detail in Chapters 2–4, we note that each of the three kinds of redox iron centers, FeS clusters, mononuclear nonheme iron, and most often iron-tetrapyrrole macrocycle variants in the heme family are found in one or more of the enzymes ([Figure 5.5](#)). Typical roles for Fe_4S_4 clusters and cytochrome b and cytochrome c forms of heme are as one-electron transfer agents. [11](#) With regard to the ubiquity of nitric oxide as intermediates noted above, the ferrous heme and ferrous-nonheme centers are high-affinity sites for metal-nitrosyl radical ligations in many of these enzymes. Three forms of structurally altered iron tetrapyrroles are utilized as redox cores for axial ligation and then: (1) reduction by six electrons *via* siroheme [12:13](#) during reduction of nitrite to NH_3 , or (2) by one electron *via* heme d1-mediated reduction of nitrite to nitric oxide by cytochrome cd1, or (3) hydroxylamine (P460 in hydroxylamine oxidoreductase) [14](#) ([Figure 5.5](#)).



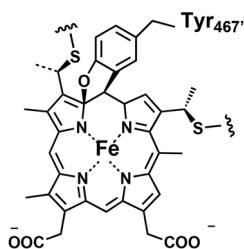
siroheme
(octacarboxy)

six-electron
nitrite reductase



heme d1

one -electron
nitrite reductase



P460 heme

hydroxylamine
oxidoreductase

Figure 5.5 Three forms of modified heme cofactors: siroheme, heme d1, heme P460.

Copper ions show up as a pair of mononuclear copper ligand sets in the first enzyme in the nitrification pathway, ammonia oxidase, where the CuA site is oxygenation catalyst. Nitrous oxide reductase, the terminal catalyst in the denitrification pathway, has six copper ions, two as a binuclear copper center serving as a one-electron shuttle while the second Cu₄S₁ tetranuclear copper cluster is the proximal electron donor to nitrous oxide (Figure 5.6).

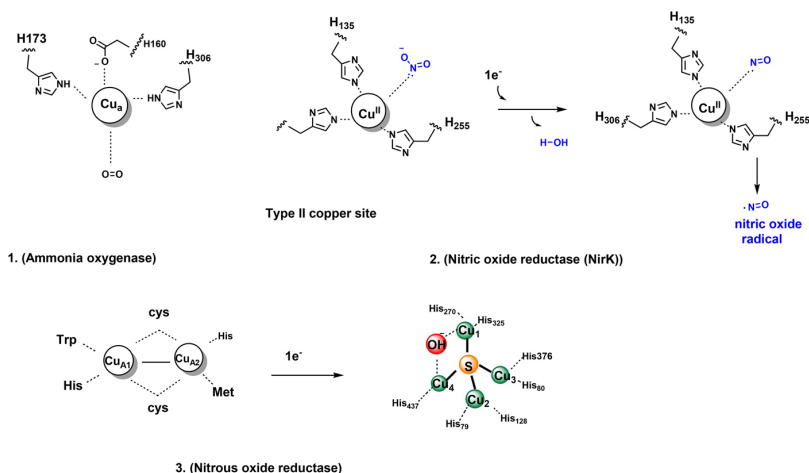


Figure 5.6 Three distinct types of redox active copper sites in the inorganic nitrogen redox pathways.

Molybdenum shows up in two enzymes, the assimilatory six-electron-reducing nitrite reductase (anaerobic pathway 1 of Figure 5.2) and in nitrogenase (Figure 5.7). The former is present as redox active molybdopterin, the latter in the M-cluster of nitrogenase where the Mo atom is not thought to be redox active. The replacement of molybdenum by vanadium in the vanadium nitrogenase is likewise not thought to be itself redox active, albeit it is a key component of the V-cluster that is the site of N₂ reduction. The two molybdenum cofactors show highly distinct roles for molybdenum in each cluster.

Nitrite reductase in
assimilatory nitrate
reduction pathway

Molybdopterin guanine dinucleotide

(6 e⁻ reduction)

Molybdenum-nitrogenase

Fe₇S₉CMo = M cluster

(6e⁻ reduction)

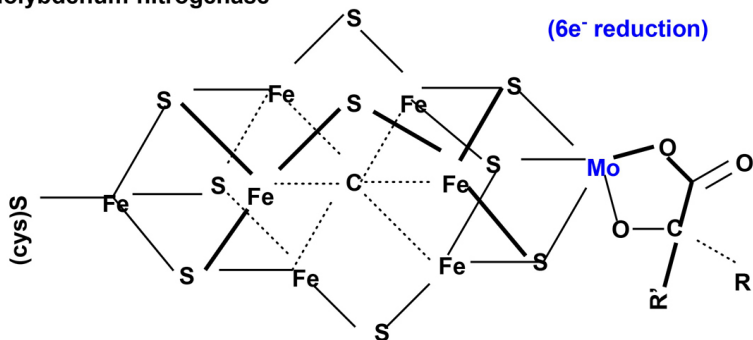


Figure 5.7 Two forms of molybdenum cofactors (FeMoCo) in nitrogenase and molybdenum guanine dinucleotide in nitrate reductase.

5.6 Turning to Organic Nitrogen Chemical Biology

As we turn our attention from inorganic nitrogen chemical functional groups, their metabolic interconnections and the routes by which plentiful but biologically inaccessible dinitrogen gas has been reductively unlocked to nucleophilic ammonia, we focus in the next section of the book on the common nitrogen functional groups in organic scaffolds. Ammonia fixation into organic scaffolds, *biological organification of ammonia*, is the universal entry point into the panoply of common nitrogen metabolites, from amino acids to amides to the nitrogen heterocycles that are the backbones of RNAs and DNAs. In most of that parallel universe of organic nitrogen chemical biology, nitrogen redox chemistry is much more muted. Ammonia and derived amines are the major nitrogen currency.

References

1. Editors Biology Dictionary, *The Path of Nitrogen through its Biogeochemical Cycle*, 2019, biologydictionary.net.
2. *Encyclopedia Britannica*, *Nitrogen Fixation*, ed. M. Petruzello, 2020, www.britannica.com.
3. S. E. McGlynn, E. Boyd, J. W. Peters and V. J. Orphan New insights into the evolutionary history of biological nitrogen fixation, *Front. Microbiol.*, 2013, **4**, 201.
4. W. G. Zumft Cell biology and molecular basis of denitrification, *Microbiol. Mol. Biol. Rev.*, 1997, **61**, 533-616.
5. S. Winogradsky Recherches sur les organims de la nitrification, *Ann. Inst. Pasteur*, 1890, **4**, 213-221.
6. M. Strous, E. Pelletier, S. Mangenot, T. Rattei, A. Lehner, M. W. Taylor, M. Horn, H. Daims, D. Bartol-Mavel, P. Wincker, V. Barbe, N. Fonknechten, D. Vallenet, B. Sgurens, C. Scheniwitz-Truong, C. Medgue, A. Coloongro, B. Snel, B. E. Dutilh, H. J. M. O. den Camp, C. van der DRift, I. Cirpus, K. T. de Pas-Schoonen, H. R. Harhangi, L. van Niftrik, M. Schmid, J. Keltjens, J. van de Voseenberg, B. Kartal, H. Meier, D. Frishman, M. A. Huynen, H.-W. Mewes, J. Weissenbach, M. S. M. Jetten, M. Wagner and D. Le Pasluier Deciphering the evolution and metabolism of an anammox bacterium from a community genome, *Nat. Rev. Immunol.*, 2006, **440**, 790-794.
7. A. Dietl, C. Frousi, W. J. Maalkcke, A. Menzel, S. de Vries, J. T. Keltjens, M. S. M. Jetten, B. Kartal and T. R. M. Barends The inner workings of the hydrazine synthase multiprotein complex, *Nature*, 2015, **527**, 394-397.
8. W. J. Maalcke, J. Reimann, S. de Vries, J. Butt, A. Dietl, N. Kip, U. Mersdorf, T. R. M. Barends, M. S. M. Jetten, J. T. Keltjens and B. Kartal Characterization of anammox hydrazine dehydrogenase, a key N₂-producing enzyme in the global nitrogen cycle, *J. Biol. Chem.*, 2016, **291**, 17077-17092.
9. D. D. Thomas, L. Reimann, J. S. Isenberg, W. Flores-Santana, C. H. Switzer, S. Donzellie, P. Hussain, C. Vecoli, N. Paolocci, S. Ambis, C. Colton, C. Harris, D. D. Roberts and D. A. S. Wink The chemical biology of nitric oxide. Implications in cellular signaling, *Free Radic. Biol. Med.*, 2008, **45**, 18-31.
10. C. Carreira, S. Pauletta and I. Moura The catalytic cycle of nitrous oxide reductase-the enzyme that catalyzes the last step of denitrification, *J. Inorg. Biochem.*, 2017, **177**, 423-434.
11. S. Bandyopadhyay, K. Chandramouli and M. K. Johnson Iron-sulfur cluster biosynthesis, *Biochem. Soc. Trans.*, 2008, **36**, 1112-1119.
12. J. M. Pennington, M. Kemp, L. McGarry, Y. Chen and E. Stroupe Siroheme synthase orients substrates for dehydrogenase and chelatase activities in a common active site, *Nat. Commun.*, 2020, **11**(1), 864.
13. S. Bali, A. Lawrence, S. A. Lobo, L. M. Saraiva, B. T. Golding, D. J. Palmet, M. J. J. Howard, S. J. Ferguson and M. J. Warren Molecular hijacking of siroheme for the synthesis of heme and d1 heme, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 18260-18265.
14. J. D. Caranto and K. Lancaster Nitric oxide is an obligate nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 817-822.

Section III

This section of the book comprises nine chapters, Chapters 6-14, that deal with different aspects of nitrogen chemical biology once ammonia has been captured or incorporated into the framework of organic metabolites. We will note that the major entry point for ammonia to cross over from inorganic to organic chemical biology involves enzymatic fixation into the scaffolds of the two common amino acids, glutamate and glutamine.

The three responsible enzymes, glutamate dehydrogenase, glutamine synthase and glutamate synthase, exemplify the two types of chemical reactions of ammonia with substrate carbonyl groups in particular and in general. They are the subject of Chapter 6.

Those reactivity modes are imine formation and amide formation. Reversible addition of ammonia (and more generally amines) may be the simplest reaction of NH_3 or RNH_2 with the bevy of carbonyl groups, aldehydes and ketones, that populate so many primary metabolites. To capture the nitrogen before the imine hydrolyzes back to amine and carbonyl components organisms adapted the strategy of reductive amination. Hydride transfer from NADH to the imine reduces the carbonyl component to the primary amine, covalently fixing ammonia (amine) to the organic scaffolds. That is the logic and essence of glutamate dehydrogenase and other reductive amination enzymes.

The second mode of ammonia capture is by reaction with carboxylates of acyl groups, e.g. glutamate in glutamine synthetase. Amides are uphill energetically from carboxylate and amine fragments and always require energy input, almost always from splitting of an ATP molecule. The chemical logic is universally for the carboxylate anion to attack one of the phosphoric anhydride linkages in the side chain of ATP to create a mixed acyl-phosphoric anhydride. In those anhydrides the substrate carbonyl is now electrophilic, capturable by ammonia (or amine, RNH_2) by nucleophilic attack. Amide bonds are central to every facet of protein biochemistry, including protein biosynthesis of the amide backbones (the amide formed between two amino acids is termed a peptide).

Once inorganic ammonia has been fixed into the organic scaffolds of glutamate and glutamine, the diaspora of amino group transfer between keto acid scaffolds occurs to populate all the amino acids in cells, from the 20 proteinogenic amino acids to the many nonproteinogenic amino acids such as ornithine, citrulline and homocysteine. We will note below that this is the role for the coenzyme form of vitamin B_6 , pyridoxal phosphate, mediating imine reduction and reoxidation chemistry in transaminase active sites.

In contrast to the constant redox changes effected in inorganic nitrogen metabolism, the great bulk of organic nitrogen chemical biology occurs without change in the formal oxidation states of the amine nitrogen. The reaction sequences of amine to amide, amide to amidine, or amide to urea to guanidine occur over and over in central metabolism and offer combinatorial sets of imine and amide chemical biology. (One

can think of amidines as imines of amides and guanidines as imines of ureido groups).

Much of organic nitrogen biology occurs in the framework of nitrogen heterocycles. Two of the side chains in proteinogenic amino acids, histidine and tryptophan, contain the imidazole and indole heterocycles, respectively. The building blocks for both RNA and DNA are the heterocyclic pyrimidine and purine rings. They present amide, amidine and guanidine groups for hydrogen bonding patterns that define the A-T and G-C pairs in the double helical strands of DNA. In keeping with the likely ancient evolutions of purines and pyrimidines, the biosynthetic routes that incorporate from two to five nitrogen atoms into those heterocycles illustrate simple amination and amidation strategies.

Two other metabolic arenas where nitrogen heterocycles bolster nature's protein catalytic inventory are, on the one hand, the B vitamins and, on the other hand, the tetrapyrrole macrocycles. Seven of the eight B vitamins (all but B₅) utilize different nitrogen heterocycles to enable central metabolic transformations. B₁ offers a thiazolium cation to enable accessible substrate carbanion chemistry; B₂ acts as a two electron-one electron step down redox conduit; B₃ in its reduced coenzyme form is the diffusible hydride donor in cells. B₆ in its aldehyde coenzyme form is synonymous with amino acid metabolism. B₇ is the carrier of one-carbon fragments at the level of CO₂, while B₉ carries one-carbon units at the levels of formate, formaldehyde and methanol, shuttling the C₁ fragments between substrates. B₁₂ is most complex of all, a cobalt containing tetrapyrrolic macrocycle, cycling between three different oxidation states of cobalt in its coenzyme roles.

The tetrapyrrole macrocycles have been termed the pigments of life. When iron is in the equatorial cavity, heme-b groups (e.g. in hemoglobin) are red-brown, while the octacarboxy siroheme is violet-purple. Magnesium-containing chlorophyll tetrapyrroles are green, the cobalt-B₁₂ is salmon pink, and the nickel-containing F430 in methanogenic bacteria is yellow.

The key nitrogen building block for all those pigments of life is an aminomethylpyrrole called porphobilinogen. We examine the enzymatic construction of that 2-aminomethylpyrrole monomer, the elimination of the side chain amine to an electrophilic azafulvene framework, and then the utilization of alternating nucleophilicity at C₆ and the electrophilicity of the C₂-exomethylene to first build linear, then macrocyclic tetrapyrroles.

All in all, the varied formation and utilization of nitrogen heterocycles, from indole alkaloids to DNA and RNA to the seven B vitamins and to the tetrapyrrolic pigments of life, are a second, remarkable nitrogen diaspora to accompany that of the amino acids. The sweep of nitrogen heterocycle chemical biology shows the many facets of nitrogen reactivity and the underlying logic of life's chemistry.

One additional aspect of Section III are the chapters on *N*-hydroxylations, to siderophores and beyond, and the generation of nitriles by *N*-oxygenation reaction manifolds. Over 3000 plants make and store cyanohydrins in protected forms and then release either cyanide or in some case isothiocyanates from stored glucosinolates as defensive responses to predation. The cyanide ion, CN⁻, among the simplest of organic nitrogen compounds, is historically a human poison by axial coordination to heme-a₃ in mitochondrial cytochrome oxidase. We examine an oxidative strategy for its formation, release and reuptake into the organic metabolite pool.

The amine *N*-oxygenations lead not only to *N*-hydroxy metabolites but also to oximes, nitroso and nitro groups. While a small subset of the universe of organic nitrogen metabolites, they raise the issue of one-electron vs. two-electron redox biology, this time in the context of organic nitrogen scaffolds. Organisms have evolved two types of redox cofactors for *N*-oxygenases. Flavoenzyme *N*-oxygenases use the cosubstrate amine as a nucleophile to attack a flavin-4a-OOH adduct to carry out net [OH⁺] transfer to the attacking nitrogen. Conversely, cytochrome P450 hemeprotein

N-oxygenases generate amine radicals and then transfer an $[\text{OH}^\bullet]$ equivalent by rebound to catalyze *N*-oxygenation.

The mammalian nitric oxide synthases are another intersection between organic and inorganic nitrogen chemical biology, at the one-electron transfer interface. The nitric oxide synthases take arginine through a net three-electron oxidation on the guanidinium side chain by way of the *N*-hydroxyarginine intermediate. In a second oxygenation step, the side chain is cleaved to citrulline and the free nitric oxide radical that then serves as stimulatory ligand for the regulatory heme group in guanylate cyclases. The nitric oxide radical is also generated by one-electron oxidations or reductions in three of the inorganic nitrogen metabolic pathways described in Section II.

The convergence of nitric oxide radical production in single-cell prokaryotes and multicellular eukaryotes indicates that one-electron oxidations at nitrogen have been a constant reactivity motif throughout evolution. Nitric oxide may have served additional metabolic roles back in the dawn of bacterial physiology.

Amination Reactions: Incorporation of Ammonia Into Amino Acid Frameworks

6.1 Enzymatic Amination to Capture Ammonia: Imines and Amides

The crossover molecule from inorganic nitrogen metabolism to organic nitrogen metabolism is NH_3 . As noted in Chapter 1, the pK_a of 9.3 for protonation of NH_3 to the NH_4^+ ammonium ion, means that at the physiological pH of 7.3, 99% of the total ammonia is present as the ammonium ion. Whereas NH_3 is a strong nucleophile on nitrogen because of the available lone pair, the nitrogen in ammonium ions is protonated and not nucleophilic. Nonetheless, when the minor NH_3 species in the reversible equilibrium reacts, it is replenished from the ammonium ion pool, so even the 1% abundance of the free ammonia form at physiological pH serves to power incorporation into organic scaffolds ([Figure 6.1](#)).

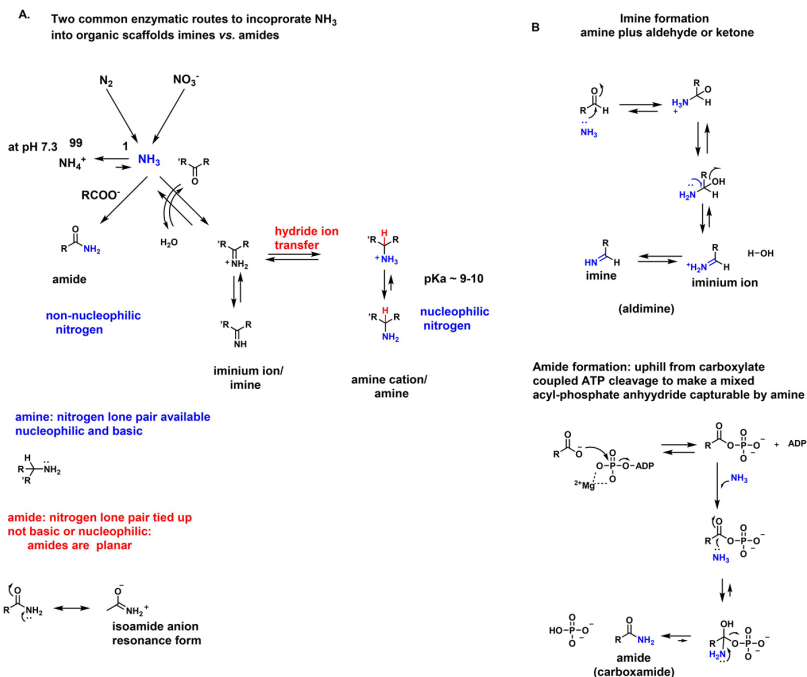


Figure 6.1 Two major routes of covalent fixation of ammonia into organic scaffolds. (A) Amine partitioning between imine formation with aldehyde and ketone carbonyls vs. amide formation between amine and carboxylates in an ATP-dependent fashion. (B). Mechanistic details for imine capture by hydride ions in reductive amination vs. acyl-phosphoric anhydride capture by amines in amide formations.

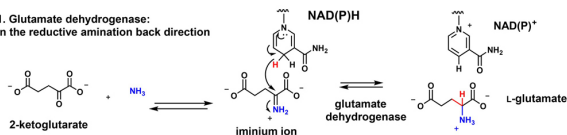
Total ammonia ($\text{NH}_3 + \text{NH}_4^+$) concentrations in human blood are in the range of 20–50 micromolar. Levels above that are toxic, especially to the brain and central nervous system. Thus, hyperammonemia (high blood ammonia) usually indicates some dysfunction of the liver (the organ that converts ammonia into urea for excretion), or possible kidney (the actual excreting organ) failure.

Ultimately, all our dietary nitrogen comes from plants through dietary intake (for non-vegetarians in part *via* animals that feed on plants). As noted in Chapter 2, the major flux for uptake of inorganic nitrogen compounds by plants is nitrate not ammonia. Assimilatory reduction *in planta* generates the nascent ammonia for the organification routes.

As depicted in Figure 6.1 the two underlying chemical steps that fix inorganic NH_3 into organic scaffolds are (1) *imine* formation with aldehydes or ketones or (2) reaction with carboxylates to form *amides*. Each of these strategies and a combination of the two are exemplified in the three enzymes that account for the bulk of NH_4^+ fixed, into the organic scaffold of the five-carbon amino acid L-glutamate (Figure 6.2A).

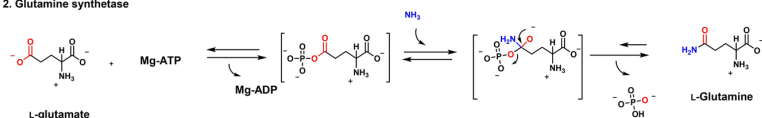
A. Capture of ammonia as reducible imine

1. Glutamate dehydrogenase:
in the reductive amination back direction



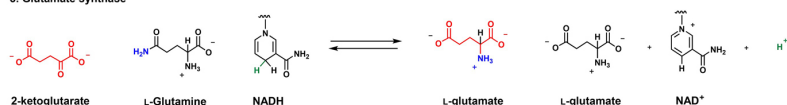
Capture of ammonia as amide

2. Glutamine synthetase



Combined amide-amine transfer logic

3. Glutamate synthase



B.

- a. Glutaminase activity on Gln generates
L-Glu and nascent "NH₃" in active site

- b. Imino-glutarate forms and is reduced by
hydride transfer. NADH acts via an active site FMN

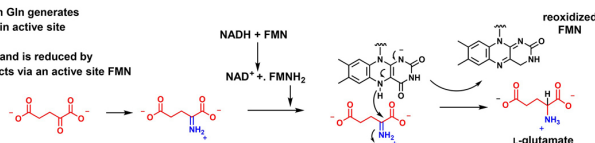


Figure 6.2 (A) Three enzymes involved in incorporation of inorganic NH_3 into the organic framework of glutamate and glutamine: reductive amination vs. amide formation; glutamine as source of nascent NH_3 by glutaminase domain action in the glutamine oxoglutarate aminotransferase (GOGAT) enzyme. (B) Combined glutaminase activity and imine reductase activity of glutamate synthase.

Before analyzing those three glutamate processing enzymes we note that imine formation between ammonia and an aldehyde or ketone carbonyl group is reversible and often unfavored at equilibrium. To capture the nitrogen permanently enzymes reduce the imine linkage (aldimine or ketimine) typically by hydride ion transfer from the reduced form of the coenzyme nicotinamide adenine dinucleotide (phosphate) [NAD(P)H(1)] (Figure 6.1). This occurs with chiral control in enzyme active site (chiral) microenvironments and typically generates an L-amino acid. The product amine from such reductive amination has a nitrogen atom still at a formal oxidation state of $\text{N} = -3$, with its lone pair available. Thus, amine group nitrogens are basic (towards protonation) and nucleophilic (towards other electrophilic partners).

In contrast, the second major route for enzymatic fixation of ammonia, exemplified by glutamine synthetase 2 in Figure 6.3 traps ammonia in a carboxamide (amide) linkage. As noted in Figure 6.1A, the lone pair in nitrogen is used to contribute to the isoamide resonance form. The C–N bond of amides has partial double bond character for that reason. Amides thus have planar geometry, compared with the tetrahedral geometry of nucleophilic amines. Amide nitrogens are not basic (not readily protonated) and tend to be, at best, weak nucleophiles. (Despite attenuation of nitrogen nucleophilicity in amides, biology does use amide and urea nitrogens in rare cases as enzymatic nucleophiles, as will be noted in Chapters 8, 10 and 16.)

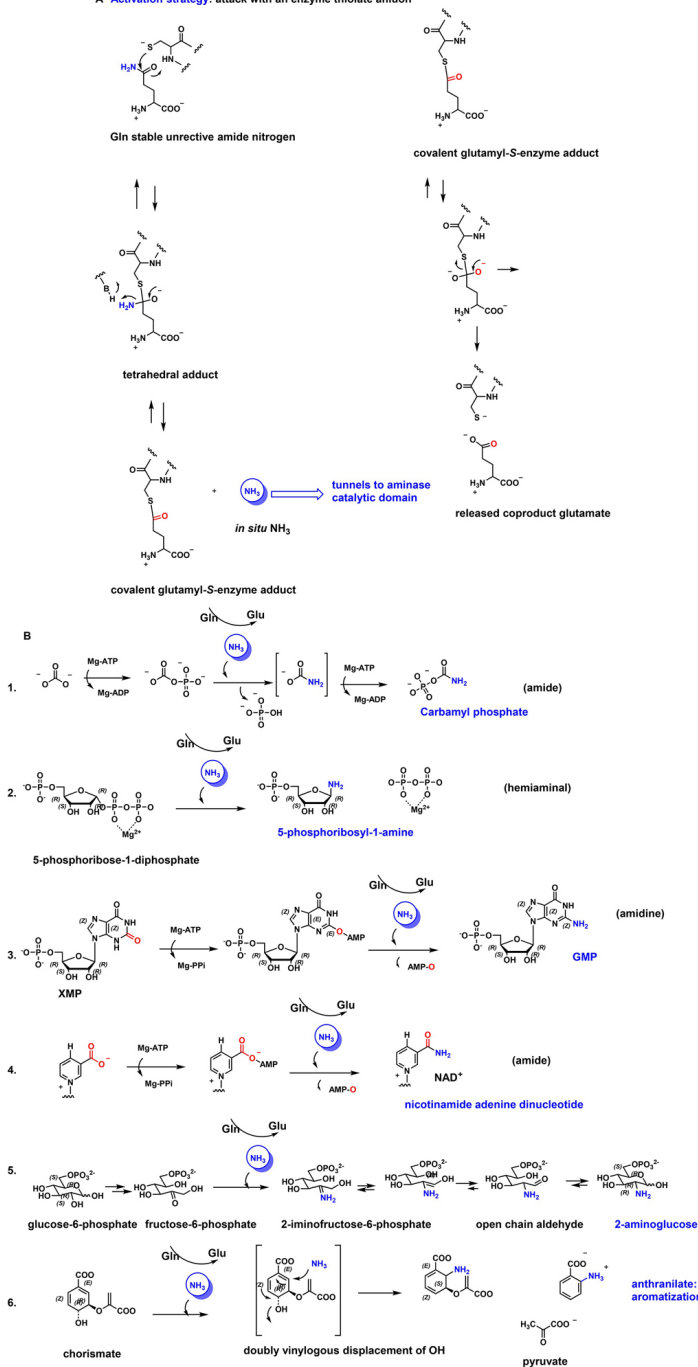


Figure 6.3 (A) Glutamine side chain carboxamido nitrogen (amide) as latent carrier of a nascent ammonia equivalent by glutaminase action, initiated by enzyme cysteine (Cys)-S⁻ attack on glutamine C₅ carbonyl of amide linkage. The glutamyl-S-enzyme covalent adduct undergoes catalyzed hydrolysis. (B) Six

examples where glutaminase domains generate nascent NH_3 utilized by companion aminase active sites: biogenesis of amide linkages in carbamyl phosphate, hemiaminal linkage in 5-*P*-ribosyl-1-amine, amidine in AMP, amide linkages in NAD^+ , amine linkages in 2-aminoglucose-6-phosphate and in anthranilate.

Figure 6.1B serves as a reminder that while imine formation is a readily reversible process between ammonia (and primary amines generally) and the aldehyde or ketone partner carbonyl, amide formation from carboxylate anions under physiological conditions is uphill energetically and *always* requires energy input for amide accumulation. The energy input is most commonly from cleavage of the diffusible cellular energy supply adenosine triphosphate (ATP). ³ The magnesium cation chelate of ATP is inevitably the substrate in ATP side chain cleavages. The Mg^{2+} acts both as a super acid to temporarily neutralize the phosphate negative charges and also to direct the regiochemistry of attack by the cosubstrate carboxylate either on P_γ or P_α in the triphosphate anhydride side chain by specific chelation geometries.

In the glutamine synthetase active site (Figures 6.1B and 6.2), the glutamate C_5 carboxylate is directed to attack P_γ of Mg -ATP to release Mg -adenosine diphosphate (ADP) and generate the bound glutamyl-5-phosphate, a mixed carboxylic-phosphoric anhydride, an *acyl phosphate*. ⁴ Acyl phosphates are thermodynamically activated for electrophilic transfer of either the acyl group at the carbonyl carbon or the phosphoryl group at the electrophilic phosphorus atom. NH_3 as the nucleophilic cosubstrate attacks the electrophilic C_5 carbonyl to give the carboxamide of product glutamine and the inorganic phosphate (Pi). The amide bond in glutamine has been driven to accumulate by the 7 kcal mol^{-1} released in hydrolysis of Mg -ATP to Mg ADP and Pi , providing a boost of approximately 10^5 on the forward equilibrium.

We will see this acyl-phosphate or related acyl-adenosine monophosphate (acyl-AMP) logic played out every time an amide bond is formed in biology. ⁴ [Sometimes the acyl-P or acyl-AMP is not captured directly by the amine component. Instead, the thiolate anion of coenzyme A (CoA) attacks to give the kinetically stable acyl CoA thioesters. These are then attacked by a substrate amine for net acyl transfer and amide formation, typically by a separate enzyme.]

6.1.1 Glutamate Dehydrogenase: Reductive Amination of Substrate Carbonyl Groups

There are three enzymes with highest flux and capacity for fixing NH_3 into organic scaffolds, all involving glutamate as participant in the reactions. The first enzyme involved in the *organification* of ammonia is the ubiquitously expressed glutamate dehydrogenase. ⁵⁻⁶ Although named for the dehydrogenation direction, the reaction is readily reversible and functions in many cellular contexts in the reverse direction to fix ammonia, as an *iminoglutarate reductase*. This enzyme catalyzes the redox-based equilibration of L-glutamate and iminoglutarate *via* hydride transfer into or out from NADH (in) or NAD^+ (out) (Figure 6.2A). In turn the iminoglutarate forms *via* rapid equilibrium with free ammonia and 2-oxoglutarate by water elimination or addition. At equilibrium the free ammonia and keto acid are favored so net trapping of ammonia only occurs when coupled to two-electron reduction of the imine to amine.

The reaction begins with binding of both substrates, ammonia or ammonium ion and 2-oxoglutarate (*aka*, α -ketoglutarate), in the active site of glutamate dehydrogenase, functioning in the reverse direction as an iminoglutarate reductase. The free base ammonia is nucleophilic at nitrogen and adds into the electrophilic carbonyl of the keto group of oxoglutarate (Figure 6.2). The initial adduct is a tetrahedral addition compound, a *carbinolamine*, that is unstable with regard to ready decomposition, in the reverse direction to starting materials or in the forward direction to the imine, in

equilibrium with the iminium ion protonated on nitrogen. The reversible addition of ammonia to a carbonyl compound to form an imine or iminium ion adduct is one of the two quintessential nitrogen-capturing reactions in organic nitrogen chemical biology (Figure 6.1). Since nitrogen is more basic than oxygen, imines are readily protonatable on nitrogen to the iminium ions, while ketones and aldehydes are not (common metabolic imine pK_a values are approximately 7, indicating a close to equal distribution of neutral imine and cationic iminium ion at neutral pH).

To proceed further along the reaction coordinate from this iminium form of glutamate ($C=N$) to the amino group of product glutamate ($C-N$) requires a two-electron reduction of the iminium ion to the corresponding amine, again a classic reaction in nitrogen chemical biology. The two-electron reduction converts a reversible carbonyl addition product (iminium ion) to the chemically stable amine group (Figure 6.1). Depending in the organism, the electron donor can be either a protein, reduced ferredoxin (one electron at a time), or, preferably, the ubiquitous small molecule cellular hydride transfer currency NAD(P)H (two electrons at a time as hydride ion). (We discuss the obligate hydride ion transfer behavior of the pyridinium group in NAD⁺ and NADP⁺ in Chapter 10.) NADH-mediated delivery of a hydride in flight to the electrophilic C_2 at one of two distinguishable faces of the iminoglutamate carbonyl in the enzyme active site yields L-glutamate.

Organification of ammonia has been achieved, converting free NH_3 to a primary amino group ($-C-NH_2$) (Figures 6.1 and 6.2A), with no change in the formal oxidation state of the nitrogen ($N = -3$). As with ammonia, the organic primary amines have pK_a values typically in the pH 9–10 range, so at physiological pH, near neutrality, the amino group in glutamate is protonated in 99 out of 100 molecules. Yet, in any nucleophilic reaction (e.g. the transaminations described in Section 6.3) it is the minor free base form that is the reactive species (Figure 6.1).

6.1.2 Glutamine Synthetase: NH_3 Organification via Amide Linkage

Glutamine synthetase is the second paradigm of capture of ammonia into an organic framework. In this case the ammonia is captured as an *amide* linkage not an amine linkage. The NH_3 equivalent being captured into an organic scaffold becomes not an amino group but part of an amide group at C_5 of the glutamate scaffold (Figure 6.2A). The group that gets amidated is one of the two carboxylates of L-glutamate.

The reaction catalyzed by glutamine synthetase is conversion of the glutamate C_5 carboxylate into the C_5 amide in glutamine. Thus, both nitrogens in glutamine, the C_2 -amino group and the C_5 amide group arise by consecutive amination or amidation reactions catalyzed by glutamate dehydrogenase and then glutamine synthetase, highlighting their centrality as metabolic entry points for ammonia organification (Figure 6.2).

As noted above and depicted in Figure 6.1B, the strategy for converting a substrate carboxylate group to a substrate amide differs from the reductive capture of an imine to an amine (Figure 6.1B). However, they both require energy input to drive amine or amide accumulation. Imines are captured by hydride transfer in reductive aminations. Each such hydride transfer spends an equivalent of the cellular electron currency NAD(P)H. Each amide formation requires cleavage of one of the side chain phosphoric anhydride linkages of cosubstrate ATP, the premier phosphoryl transfer currency in cells. NH_3 covalent incorporation by these two most common routes cost cells energy to drive metabolic fixation into organic metabolite scaffolds.

The carboxyl groups of substrates at physiological pH values are all ionized. They are oxygen nucleophiles not electrophiles so attack by the lone pair electrons in nucleophilic NH_3 is repelled electrostatically. This holds for all carboxylate to amide

transformations in biological metabolism (and is the reason energy input is required to convert the ground state nucleophilic carboxylate anion to an electrophilic derivative).

The general solution 4 is to convert the carboxylate ion from a nucleophile to an electrophile so that it can be attacked by NH_3 . The nucleophilic C_5 carboxylate of glutamate attacks the γ -phosphoryl group of Mg-ATP in the glutamine synthetase active site (Figure 6.2A). That creates Mg-ADP and the C_5 -glutamyl-phosphate (an acyl phosphate = a mixed acyl-phosphoric anhydride, doubly activated for either electrophilic acyl transfer or electrophilic phosphoryl transfer). The NH_3 cosubstrate can behave as nucleophile towards the now electrophilic C_5 of glutamyl phosphate. The resultant tetrahedral adduct can partition backwards where it came from or forward to liberate inorganic phosphate and form the amide bond at C_5 of glutamine.

Glutamine synthetase in bacteria is a key enzyme in assimilation of NH_3 and highly regulated. The bacterial Gln synthetase is famously a dodecahedron of twelve subunits stacked as two hexamers, each of which is subject to a bevy of regulatory controls, including covalent adenylation of a lysine residue, discussed in Chapter 16.

In principle, this reaction sequence could be run in reverse and, indeed in discussion of the third enzyme as well as more generally in Section 6.2 we note that glutamine serves as a carrier of a latent ammonia group, available on demand in several enzyme active sites. When ammonia is covalently tied up as the C_5 amide in l-glutamine , a readily diffusible amino acid, the nucleophilic (Figure 6.3) reactivity and potential toxicity of free ammonia has been hidden as the unreactive amide group. The principle is revealed in the strategy of the next enzyme, glutamine oxoglutarate aminotransferase.

6.1.3 Glutamine Oxoglutarate Aminotransferase

The third aminating enzyme, glutamine oxoglutarate aminotransferase (GOGAT) 7 brings the principles of the above two enzymes, glutamate dehydrogenase and glutamine synthetase, into congruence (Figure 6.2A). The chemistry involves activating the amide nitrogen of glutamine as a source of nascent ammonia, then capturing it for carbonyl addition to the ketone of oxoglutarate. The reductive amination in the second half reaction is the essence of the glutamate dehydrogenase strategy (Figure 6.2).

GOGAT is restricted in distribution to cyanobacteria and photosynthetic eukaryotes, e.g. plants 7 (where it is located in chloroplasts) but is revelatory in its catalytic logic. GOGAT is sometimes termed glutamate synthase to reflect its physiological role.

The use of glutamine as a circulating carrier of latent ammonia, released in a target enzyme's active site, is the first half reaction of GOGAT. To activate the resonance stabilized C_5 amide of glutamine the enzyme adds an active site cysteine thiolate side chain (the most powerful common nucleophile in the biologic armamentarium) into the amide carbonyl group (Figure 6.3A). The resultant tetrahedral adduct is an enzyme-tethered *thiohemiaminal*. Partition in the forward direction yields a covalent glutamyl-S-enzyme (a thermodynamically activated thioester) and the nascent NH_3 . To get ready for a subsequent turnover, the GOGAT enzyme catalyzes water addition into the thioester carbonyl to release free glutamate.

Meanwhile, it is the nascent NH_3 , its nucleophilicity now uncovered, that is the point of the first half reaction. Before it can be released, unleashed wastefully into solution, the NH_3 reacts with the bound oxoglutarate ketone carbonyl and forms the now familiar iminium adduct. Reduction by NADH or in some cases by the dihydro form of the protein ferredoxin converts the C_2 -imine to the C_2 amine of l-glutamate (Figure 6.2). The first half reaction of GOGAT is a glutaminase reaction, the second is a reductive amination. 8 These are two major features of the interface between inorganic ammonia and organic amine and amide frameworks as noted in Figure 6.1.

6.2 Glutamine as Carrier of Latent Ammonia as a General Strategy in Enzymatic Aminations

We will note in the next chapter that multicellular organisms go to great lengths to keep the concentrations of cellular ammonia below toxic levels. The nucleophilicity of ammonia is problematic, so organisms from bacteria to humans have generalized the use of glutamine as a latent carrier for ammonia in a non-nucleophilic form. The ammonia can be liberated in nascent form using the glutaminase logic evinced above in the first half reaction of the GOGAT enzyme and generalized in [Figure 6.3A](#).

Glutaminases, often embedded as separate domains within a family of enzymes known formally (but mechanistically misleadingly) as amidotransferases, [9](#) are found in at least 16 enzymes that liberate equivalents of nascent ammonia *in situ* in a glutaminase active site *via* the mechanism shown in [Figure 6.3A](#). The glutaminase can be a separate domain from the catalytic amination domain or it can be a separate subunit. [10](#) Whether intraprotein or interprotein, typically the glutaminase–aminase pairs have tunnels between the domains or subunits for nascent ammonia to diffuse to the amination active site without exchanging with solution and undergoing dilution and release of this potentially toxic nucleophile [11](#) ([Figure 6.4](#)).

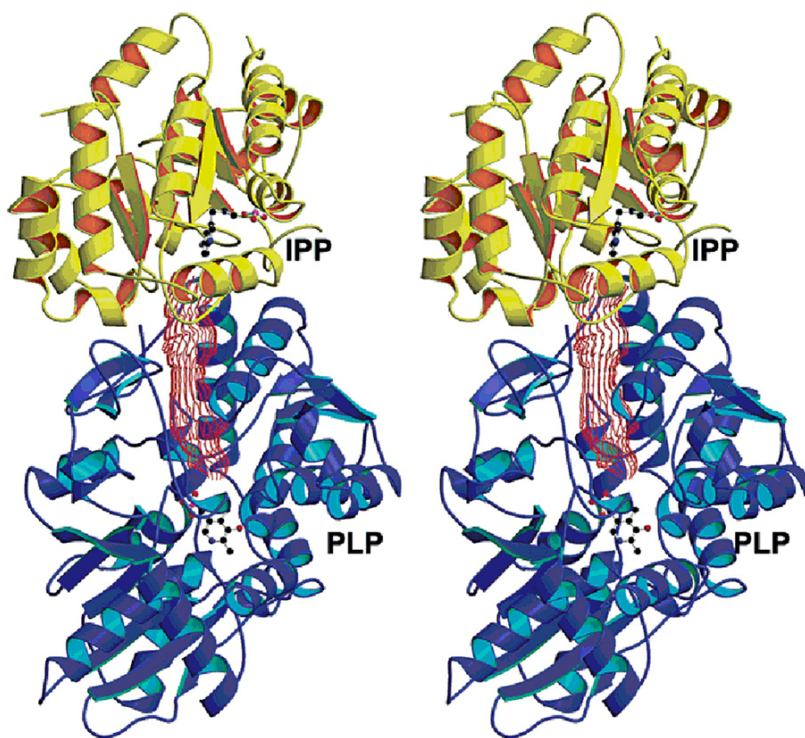


Figure 6.4 Glutaminase-containing amidotransferases typically contain two subsites, one for the glutaminase and one for the specific aminase site. Nascent NH_3 released by glutaminase action diffuses through internal tunnels to the aminase active site without getting free into solution. The diagram shows the

This 16 member enzyme family has been termed the *amidotransferase family* historically to denote the use of glutamine as amino donor. [9](#) While technically correct, the term is mechanistically misleading. It is not an amido group that is being transferred. The non-nucleophilic amido group is being hydrolyzed *in situ*, to release nucleophilic NH_3 as the proximate amino donor ([Figure 6.3A](#)). Mechanistically, the transfers are amino transfers not amidotransfers. These biosynthetic reactions are irreversible in cells by virtue of the fact that net hydrolysis of glutamine to glutamate (amide to acid anion) is physiologically and thermodynamically irreversible. This is a useful energetic parameter to drive a biosynthetic equilibrium in the forward direction while averting the toxicity of saturating solution levels of NH_3 for each “amido” transferase.

The cosubstrates that get aminated vary in structure and location in biosynthetic pathways. The simplest is carried out by the enzyme carbamoyl-phosphate synthase, [12](#) creating the mixed anhydride between carbamic acid and phosphoric acid ([Figure 6.3B](#), reaction 1). We will note its downstream utilization as nitrogen carbamylating agent, both in urea biosynthesis (Chapter 7) and in the pyrimidine biosynthetic pathway for both RNA and DNA building blocks (Chapter 8).

Several glutaminase domains or subunits are involved in amino acid biosynthetic pathways – leading to asparagine, histidine, and tryptophan. In both purine ring assembly at the earliest stage to generate β -1-amino-ribose-5-phosphate ([Figure 6.3B](#), reaction 2) and at the very end where guanosine monophosphate (GMP) is made from inosine monophosphate (IMP) ([Figure 6.3B](#), reaction 3), glutamine provides the nascent ammonia equivalent. [Analogously, at the late stage of pyrimidine maturation, cytidine triphosphate (CTP) is made from uridine triphosphate (UTP) by CTP synthase using glutamine as a donor of nascent ammonia (not shown).] The final stage in assembly of the redox active coenzyme NAD^+ is schematized in reaction 4 of [Figure 6.3B](#) while the creation of 2-aminoglucose for arrays of posttranslational protein N-glycosylations (see Chapter 16) is shown in reaction 5. [9](#)

Two of these five examples use the nascent *in situ* NH_3 for *amide* formation (carbamyl-phosphate and NAD^+). Meanwhile, the phosphoribosyl diphosphate amination creates the *hemiaminal* linkage at C_1 in 5-phosphoribosyl-1-amine on which the purine ring is biosynthetically assembled. The 2-amino glucose reaction involves initial imine formation at C_2 of fructose-6-phosphate followed by a 1,2-hydroxycarbonyl rearrangement that traps the initial imine of the fructose scaffold as the C_2 amine of the glucose scaffold. The IMP to GMP amination functionally provides one of the four nucleobases for both RNA and DNA biosynthesis. Chemically, it is conversion of an amide to an *amidine*, introducing an organic nitrogen functionality crucial to hydrogen bonding schemes in DNA (see Chapter 8 on amidines).

The final example in [Figure 6.3B](#) of the use of glutamine-derived nascent NH_3 is in anthranilate synthase. The product anthranilate, (aka *ortho*-aminobenzoate) is a building block for tryptophan and is an *arylated amine*. Its formation derives from the common dihydroaromatic precursor metabolite chorismate. [13](#) The reaction proceeds in two steps: first a doubly vinylogous displacement of the $\text{C}_4\text{-OH}$ of chorismate by attack of nascent NH_3 at C_2 . Then, the aminated intermediate undergoes an aromatizing elimination of the elements of pyruvate across C_5 and C_6 to yield anthranilate. The corresponding *para*-aminobenzoate regioisomer, a building block for vitamin B_9 folic acid (Chapter 10), is assembled by analogous enzymology from a glutamine-derived active site ammonia addition to chorismate [13](#) (not shown).

The utility of glutamine as a “just in time” nonnucleophilic amide donor of nascent

nucleophilic ammonia at a target enzyme site has radiated into many biosynthetic aminations across multiple metabolic pathways. The strategy has likewise led to a ramification of distinct nitrogen functional groups across various chemical scaffolds.

6.2.1 Asparagine and Aspartate vs. Glutamate and Glutamine

The closest homologs to glutamate and glutamine among common amino acids in primary metabolism are the congeneric aspartate and asparagine, respectively (Figure 6.5A). Asparagine synthetase has a comparable task to glutamine synthetase in the need to activate the substrate carboxylate, here regiospecifically the β -carboxylate of aspartate. It reveals parallel logic but with two differences in execution strategy.

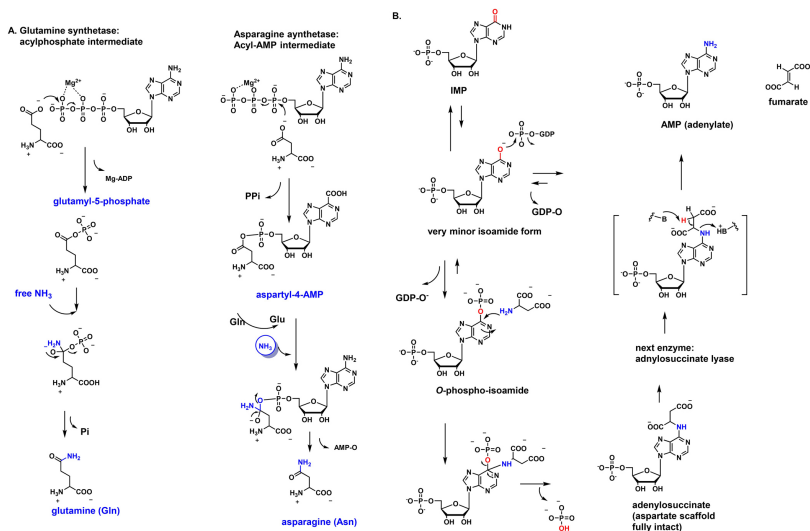


Figure 6.5(A) Parallel strategies for glutamine and asparagine side chain amide formations: two modes of ATP side chain fragmentation. (B) Aspartate as amino donor by covalent addition–elimination *via* the adenylosuccinate synthase–adenylosuccinate lyase pair of enzymes, releasing the four-carbon scaffold as fumarate.

Mg-ATP is again the thermodynamically activated cosubstrate utilized, but instead of attack by one of the carboxylate oxygens on P_γ of ATP, asparagine synthetase directs attack at P_α , generating a distinct mixed acyl–phosphoric anhydride intermediate. Now, it is tightly bound aspartyl-AMP, rather than aspartyl-phosphate, that is the activated species. Second, the donor of the attacking cosubstrate ammonia is not free ammonia from solution. Instead, it derives from action of a glutaminase domain, with the logic of the glutamine synthetase catalyst (one of the 16 enzyme cases noted above). Glutamine is hydrolyzed by the usual way of a glutamyl–S–Cys-enzyme adduct and the nascent NH_3 travels down an internal protein tunnel to where aspartyl-AMP is bound. ¹⁴ The *in situ* NH_3 attacks the activated β -carbonyl group of aspartyl-AMP. The result is net release of free AMP and formation of asparagine.

In contrast to the 16 enzymes that use glutamine as a latent source of NH_3 for biosynthetic aminations as discussed above, asparagine is *not* used as a source of latent ammonia. Furthermore, there is no NAD(P)-dependent aspartate dehydrogenase or

iminoxaloacetate reductase comparable to the glutamate dehydrogenase that acts as an NH_3 fixation catalyst (refer back to [Figure 6.2](#)). Therefore, glutamate and glutamine remain the main pair of entry points for *organification* of NH_3 by reductive imine or amide capture.

On the other hand, aspartate itself is utilized by organisms, including humans, as an amino group donor, but in a quite distinct chemical context. Aspartate undergoes enzyme-catalyzed adduct formation by an enzyme in the biosynthetic pathway to AMP *via* an adenylosuccinate intermediate [15](#) ([Figure 6.5B](#)), hundreds of millions of times per cell cycle. The deprotonated amino group of aspartate is nucleophilic. It adds into the C_6 position of IMP only after the rare isoamide form of IMP has attacked cosubstrate GTP at P_γ to yield the phosphorylated isoamide. Aspartate amino group addition to C_4 is followed by expulsion of inorganic phosphate, containing the original oxygen that had been the C_6 carbonyl of IMP. The resultant adduct, adenylosuccinate, has a new $\text{N}-\text{C}_6$ bond to IMP, but the aspartate framework is still intact.

Completion of AMP biosynthesis, by a net aspartate *addition* and subsequent cleavage of the aspartate- C_2-N *via* fumarate *elimination* is catalyzed by the next enzyme in the purine pathway, adenylosuccinate lyase. [16](#) As pictured in [Figure 6.5B](#), the lyase is believed to initiate catalysis by abstraction of one of the two prochiral hydrogens at C_3 of aspartate. The electron pair on C_3 acts as a neighboring group to create the *trans* olefinic C_3-C_2 double bond as the C_2-N bond cleaves. The resultant products are fumarate and AMP (known historically as adenylic acid).

In a mechanistically analogous second such transformation, aspartate gives rise to argininosuccinate as the key precursor to urea for the safe excretion of ammonia equivalents, to be examined in Chapter 7. The use of the amino group of aspartate as a net donor of an “ NH_3 ” in metabolic pathways is rare compared with the glutaminase mechanism. Hydrolysis of the glutamine carboxamide generates nascent NH_3 sequestered and set to tunnel in target enzyme active sites. The aspartate addition-elimination two step mechanism avoids free “ NH_3 ” completely. We return to this amino donor dichotomy in Chapter 9 on nitrogen heterocycles and purine biosynthesis strategies.

6.3 Moving Amino Groups Around on Organic Scaffolds

We have noted above that the ammonia-derived carboxamide nitrogen of glutamine is mobilized for some sixteen *in situ* hydrolyses to power aminations, including aspartate to asparagine. However, a separate logic and strategy is employed for shuttling the amino group of glutamate to the other 18 proteinogenic amino acids and many nonproteinogenic amino acids as well. A family of approximately 30–40 enzymes in humans known to contain one tightly bound equivalent of the coenzyme form of vitamin B_6 , pyridoxal-phosphate (PLP), are known as transaminases. [17](#) As depicted in [Figure 6.7](#), the transaminases shuttle the amino group between a donor amino acid and an α -keto acid recipient.

The pyridoxal-phosphate is the essential cofactor lowering the energy barriers for shuttling the amino group between amino acid donor and keto acid recipients. The key functional group is the exocyclic aldehyde appended to the pyridine core. [1](#) There are no aldehyde groups in the side chains of the 20 proteinogenic amino acids, so this is a novel functional group addition to nature's catalytic inventory. The concentration of free PLP in cells is almost negligible. Each transaminase and other subclasses of PLP enzymes bind the aldehydic cofactor *via* a side chain lysine residue C_6-NH_2 group. The first step in every transaminase is transaldimination between the side chain lysine

residue in the active site and the amino acid substrate. For simplicity we have omitted the starting and ending transaldimination steps from the figures and mechanistic discussion below.

6.3.1 Transaminase Logic

Five general facets of transaminase catalysis are summarized in [Figure 6.6](#). (1) PLP is an effective electron sink for delocalizing and thereby stabilizing aminoacyl- C_2 carbanion equivalents. (2) Transaminases are *redox catalysts*, performing two half reactions. In the first half reaction the amino acid substrate is oxidized by two electrons as the PLP coenzyme is reduced by two electrons [to the pyridoxamine-phosphate form (PNP)]. This is reversed in the second half reaction as keto acid-2 is reductively aminated to amino acid-2 while PNP is reoxidized to PLP. (3) The NH_3 group that is in transit between amino acid and keto acid substrates is stored temporarily but covalently as pyridoxamine-phosphate. (4) In keeping with the general avoidance of toxic free ammonia in cells, there is no free NH_3 in transaminase catalytic cycles. The “ NH_3 ” in transit is always covalently tethered to the pyridoxal framework. (5) Transaminase catalytic logic involves reversible formation of initial imine (*aldimine*) formation between amino acid-1, then the redox step that creates ketoacid-1-imine to PNP ([Figure 6.6B](#)). Release of keto acid-1 allows keto acid-2 to bind in *ketimine* linkage. Two-electron reductive amination involves generation of amino acid-2 imine to PLP ([Figure 6.6B](#)).

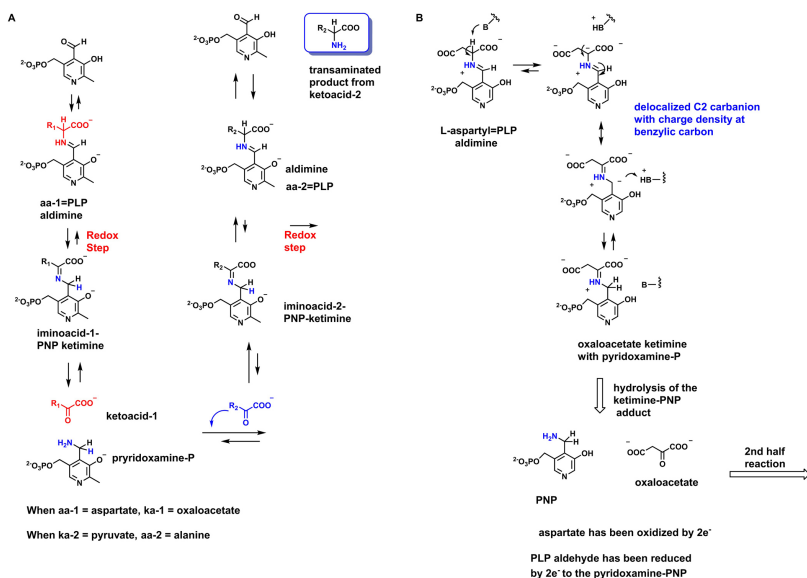


Figure 6.6 (A) Transamination by PLP-enzymes: moving the itinerant amine group between keto acid scaffolds with PNP intermediacy to avoid free NH_3 . (B) The redox step in the first half reaction involves protonation of the C_2 delocalized carbanion at the benzylic carbon of the B_6 coenzyme: a net 1,3-proton shift.

Different transaminases show distinct ranges of specificity for amino acid donor and keto acid receptor for the itinerant amino group. [Figure 6.7](#) shows three sets of (hypothetical) transaminases. The first uses L-alanine as an amino donor and three ketoacid acceptors are depicted, generating L-aspartate, L-valine, and L-glutamate. In the

reverse direction L-glutamate and pyruvate will give the observed L-alanine. Starting from L-glutamate one can get to L-alanine and start the amino group dispersion to populate other amino acid metabolites, mediated by different members of the transaminase family. In [Figure 6.7B](#) glutamate is the amino donor to the three indicated common α -ketoacids to yield L-alanine, L-aspartate and L-methionine. Analogously, a third transaminase uses L-aspartate as amino donor to three aromatic α ketoacid metabolites to give L-isomers of phenylalanine, tyrosine and tryptophan. The indicated three amino acids L-alanine, L-glutamate and L-aspartate are the most common amino donors in the transaminase family-mediated diaspora.

Moving Amino Groups Between 2-Keto Acid Scaffolds: Transaminases

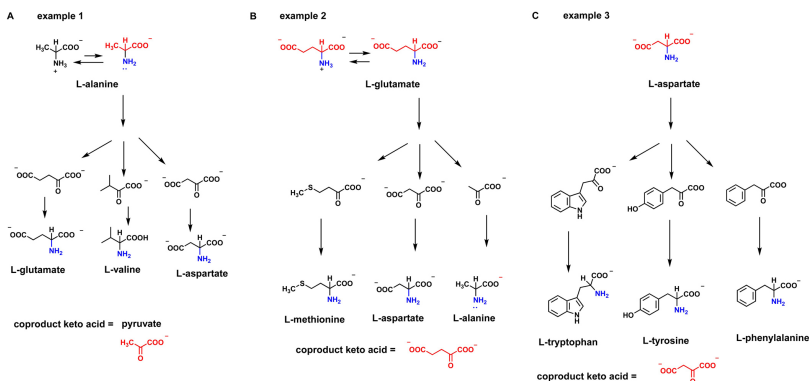


Figure 6.7 Diaspora of amino groups shuttled between amino acid and keto acid frameworks. (A) L-Alanine as amino donor to three keto acids generates L-glutamate, L-valine and L-aspartate; (B) L-glutamate as amino donor to three keto acids generates L-methionine, L-aspartate and L-alanine; (C) L-aspartate as amino donor to three keto acids generates L-tryptophan, L-tyrosine and L-phenylalanine.

Transaminases are fully reversible catalysts so the amino acids can be precursors of the keto acids in the reverse direction. This is relevant when a large protein load arrives from dietary intake: breakdown of the proteins to the constituent twenty proteinogenic amino acids causes an acute nitrogen load. The excess nitrogen can accumulate in aspartate by a series of transaminase shuttles and the aspartate nitrogen is then converted to urea for excretion as noted in Chapter 7.

Among other metabolites are two aldehyde-acid metabolites that can function as transaminase cosubstrates. Glyoxylate on reductive amination, say by aspartate, gives the two-carbon amino acid glycine as the aspartate framework is converted to oxaloacetate. Finally, succinic semialdehyde (half aldehyde, half acid) is reductively aminated to gamma-aminobutyrate (GABA), an inhibitory neurotransmitter in the central nervous system (CNS). A typical amino donor cosubstrate would be glutamate, undergoing net oxidation to 2-oxoglutarate.

6.3.1.1 Transaminases: Two Half Reactions

Transaminases undertake two half reactions that can be written as an oxidative half reaction and then a reductive half reaction. [17](#) In the oxidative half reaction the substrate, PLP aldimine, is converted to a nascent product, PNP ketimine. The amino acid framework in this case is oxidized to the imine. The pyridoxal aldehydic carbon has been reduced by the equivalent two electrons to the pyridoxamine (PNP)

component of that ketimine adduct (Figure 6.6B).

This oxidative half reaction is initiated by proton transfer chemistry in the enzyme active site. Catalysis starts with an enzymic active site base cleaving the Asp C₂-H bond as a proton abstraction. The energy of forming the resultant Asp C₂ carbanion has been substantially lowered by the ability to delocalize that carbanion through the PLP pyridine ring system as shown. Indeed, it is the functioning of PLP as an electron sink to delocalize amino acid substrate C₂ carbanions that makes PLP synonymous with amino acid chemical biology in dozens of metabolic pathways.

The fate of the delocalized Asp-C₂ = PLP carbanion depends on the placement of protonated side chain bases (conjugate acids) in a given enzyme active site. ¹⁷ In all transaminases there is another -BH⁺ group (-B₂H⁺) oriented in proximity not to C₂ of the Asp substrate but to the exocyclic aldehydic carbon (the benzylic carbon) of the PLP moiety. One of the resonance contributors to the delocalized carbanion in the Asp = PLP system has carbanion density at that PLP carbon (Figure 6.6B). Regiospecific proton delivery to that site has the following consequences. It traps the pyridoxal moiety as a two-electron reduced pyridoxamine (PNP). The four-carbon substrate is in turn connected now to the PNP as a ketimine.

The first half reaction is almost complete. There has been an internal two-electron redox reaction, mediated by two regiospecific proton transfers to convert Asp = PLP aldimine to oxaloacetate-PNP ketimine. At this juncture the oxaloacetate ketimine can be attacked by a water molecule in the active site, with net release of oxaloacetate and leaving the enzyme in the two electron-reduced pyridoxamine-phosphate state. This snapshot emphasizes the redox nature of the first half reaction.

6.3.1.2 The Second Half Reaction: Reoxidation of the Enzyme-PNP to Enzyme-PLP

The PNP-form of the transaminase is incompetent to undergo another catalytic turnover cycle with aspartate as substrate. On the other hand, it could go backwards with oxaloacetate all the way to aspartate. That would constitute an invisible virtual reaction. However, if instead a molecule of pyruvate, an abundant keto acid metabolite in cells, could bind in the active site, it could go on to form an imine with the PNP cofactor. This iminium ion, between the keto acid pyruvate and PNP, is a ketimine. This is the start of the second, reoxidative half reaction of this transaminase.

Catalysis involves abstraction of one of the two CH₂NH = pyridoxamine prochiral hydrogens as a proton; this would be the microscopic reverse of the last step of the first half reaction. It would create again the stabilized, delocalized carbanion, with electron density at that pyridoxal benzylic carbon but also at C₂ of the tethered iminopyruvyl moiety. Proton delivery to that C₂ would create the L-alanyl = PLP imine. Net transaldimination by re-attack of the active site lysine-NH₂ side chain liberates L-alanine and restores the starting enzyme-NH = PLP adduct ready for the next catalytic cycle.

The reoxidative half reaction has converted the PNP form of the B₆ coenzyme back to the aldehyde of PLP. It has reductively transferred the amine group to pyruvate. Overall, one can see that the two half reactions had the net effect of moving the amine group from the C₄ scaffold of L-aspartate to the C₃ scaffold of L-alanine. The -NH₂ group in flight between the two organic scaffolds is *never free as NH₃*. Instead, it is tied up covalently to the pyridoxamine cofactor in its two-electron reduced intermediate state. Although the reaction stoichiometry of transaminases does not reveal redox chemistry it is the underlying strategy for all transaminases.

The PLP coenzyme can and does mediate amino acid carbanion chemistry on all four substituents at C₂ of different amino acid substrates as well as at C₃ and C₄ of some substituted amino acids. We will return to its carbanion stabilizing properties briefly in

the next section on certain amino acid deaminase reactions that do release free ammonia as product.

6.4 Reactions that Release Ammonia

6.4.1 Amino Acid Deaminases and Ammonia Lyases

The functional reverse of the three glutamate -utilizing enzymes described in Sections 6.1.1–6.1.3 that covalently incorporate free NH_3 into the glutamate or glutamine organic scaffolds are enzymes that release free ammonia from organic scaffolds as part of their physiological functions. These fall into a broad category of *deaminases*, sometimes known as ammonia lyases. The lyase terminology means cleavage of a substrate, in these cases of C– NH_2 carbon–nitrogen bonds to release free ammonia.

In a physiological sense the ammonia lyases–deaminases balance the actions of the three glutamate enzymes that fix NH_3 into the glutamate and glutamine frameworks. These balancing fluxes of catalytic NH_3 fixation and release give a sense of the dynamic flux between enzymatic incorporation into or disgorgement of NH_3 from organic metabolite frameworks.

A recent review [18](#) lists 22 ammonia lyase enzymes that function in discrete cellular contexts. Most of them involve amino acid substrates that undergo catalyzed deaminations to release NH_3 . However, in contrast to the many transaminases noted in the previous section that employ a common PLP-based mechanism for all members of that enzyme class, there are multiple mechanisms operant in the ammonia lyases ([Figure 6.8](#)).

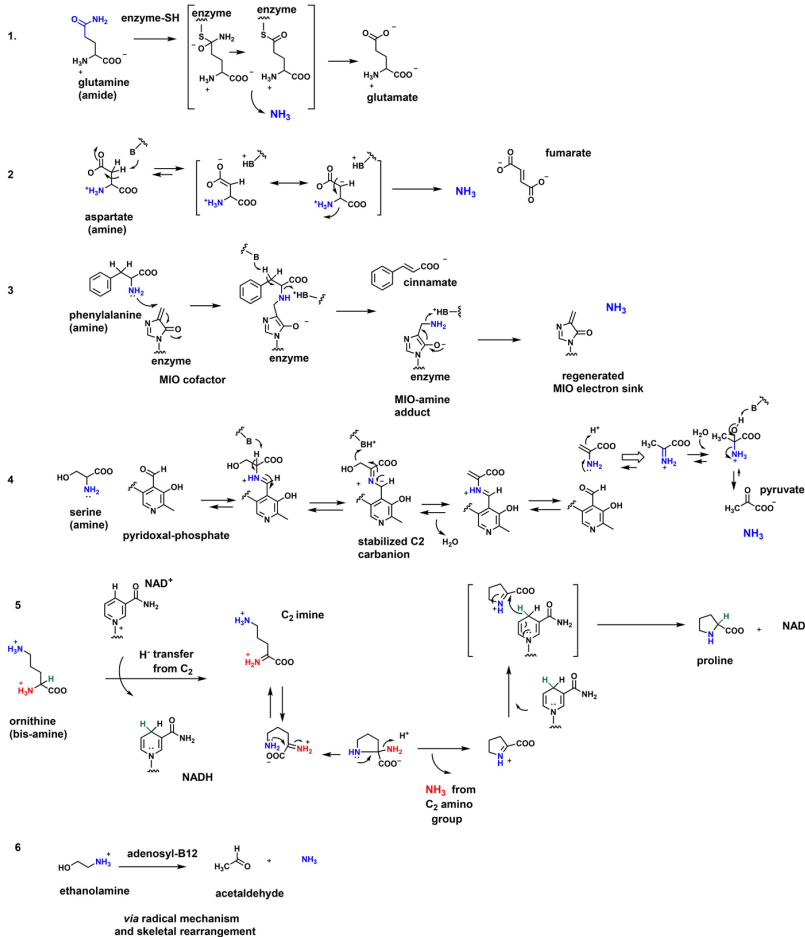


Figure 6.8 Ammonia lyases employ diverse mechanisms to eliminate ammonia while generating a C=C double bond in the coproduct. Five amino acid lyases, for glutamine, aspartate, phenylalanine, serine and ornithine, show the range of catalytic strategies that have evolved for these deaminases. The sixth example shows action of a B_{12} -dependent mutase on the amine substrate ethanolamine.

The 16 glutaminases noted in Section 6.2 that deliver nascent NH_3 only at the active sites of target enzymes are *cryptic* deaminases in the sense that the NH_3 never (hardly ever) leaks into free solution as product. Instead, that nascent NH_3 is used for amination of cosubstrates as summarized in Figure 6.3B. This cryptic action of glutaminases domains as deaminases are noted in line 1 of Figure 6.8.

In line 2 of Figure 6.8, aspartate ammonia lyase is also a coenzyme-independent catalyst. A proposed mechanism involves carbanion generation and stabilization through an *aci*-carboxylate form that then ejects the NH_3 substituent to give the *trans*-olefinic diacid fumarate as coproduct. This is similar to the reaction mechanism proposed above for fumarate generation by adenylosuccinate lyase. 16

In contrast, four other sets of ammonia lyases employ coenzymes. L-Phenylalanine (line 3, Figure 6.8) and histidine in plants are deaminated to the respective olefins by an unusual methylene imidazolone (MIO) cofactor that arises from three residues in the enzyme during maturation to the active catalyst. The MIO is an electron sink that

allows for C–N bond cleavage in the phenylalanine framework, creating cinnamic acid that is the source of thousands of phenylpropanoid secondary plant metabolites. ¹⁹

In contrast, L-threonine and L-serine (line 4 of [Figure 6.8](#)) undergo more or less conventional deamination by PLP-dependent deaminases, via stabilized carbanion mechanisms to 2-ketobutyrate and pyruvate, respectively. ¹ So do 2,3-diaminopropionate and 2-amino-6-phospho-D-glucose. Unlike the PLP-containing transaminases, in the PLP-dependent deaminases the stabilized substrate C₂-carbanion is not used to protonate the benzylic carbon of the pyridoxal but rather as a neighboring group to expel the β-OH substituent from serine or threonine. For example, this net elimination of the elements of water across C₂ and C₃ of the serine framework generates a three carbon en amino-PLP adduct that is ultimately precursor to pyruvate and free ammonia.

The linear C₅ di-amino acid ornithine is precursor to the proteinogenic proline, the only secondary amine used in protein biosynthesis ²⁰ ([Figure 6.8](#), line 5). The conversion of linear ornithine to cyclic proline involves loss of one of the two amino nitrogens and also requires a tightly bound equivalent of the oxidized form of the hydride transfer coenzyme NAD⁺. The chemical logic is for C₂ of bound ornithine to be oxidized from the amine to the imine by hydride transfer, creating a two electron-reduced molecule, bound NADH. This an example of the reverse of the paradigm in [Figure 6.1](#) for fixation of NH₃ into substrate carbonyl groups by imine formation followed by two-electron reduction to the amine. Here, the purpose is the opposite: to expel NH₃ by amine oxidation and resultant imine hydrolysis.

The hydride transfer from C₂ of bound ornithine to bound NAD⁺ creates the C₂ electrophilic imine that is capturable by intramolecular attack of the C₅–NH₂ acting as a nucleophile. The resulting tetrahedral adduct has two nitrogens (a cyclic amina), [Figure 6.8](#), line 5). Decomposition in the forward direction, expels what had been the C₂-nitrogen as ammonia, producing the cyclic five membered ring iminium ion. Now the final step is hydride transfer back from NADH to set the oxidation state and stereochemistry of the proteinogenic cyclic amino acid L-proline. The enzyme is back in its resting NAD⁺ oxidation state, ready for the next catalytic cycle.

One more deaminase variant (line 6 of [Figure 6.8](#)) is in the oxidative metabolism of ethanolamine (derived from serine by metabolic decarboxylation) by some bacteria that utilize coenzyme B₁₂ (one of the two biologically active forms of vitamin B₁₂, see Chapter 10). These bacteria convert ethanolamine to propionaldehyde and free ammonia. This turns out to be a classic example of B₁₂-initiated radical chemistry on the ethanolamine substrate. We examine B₁₂ in the context of vitamin-derived nitrogen heterocycles in Chapter 10 and give an overview of its biosynthesis in Chapter 11.

It is clear that a variety of reaction mechanisms have evolved to carry out cleavage of the C–N bond of amine-containing substrates. In the glutaminase examples it is the nascent NH₃ that is the valuable product, not the glutamate. In the other deaminases of [Figure 6.8](#), it is the NH₃ that is essentially a waste product. The altered carbon scaffolds of the coproducts are the valuable metabolites.

6.4.2 Deamination by Amino Acid Dehydrogenases and Oxidases

One more path to deamination of amino acids to the corresponding 2-ketoacids and free NH₃ is supplied by the action of various dehydrogenases and oxidases in cells. Dehydrogenases and oxidases are often distinguished definitionally by whether the two electrons removed from the amino acid as it is oxidized are transferred to molecular oxygen (oxidases) or some other electron acceptor (dehydrogenases). Typical dehydrogenase cosubstrates are oxidized NAD⁺ that undergoes two-electron reduction to NADH as the imino-acid coproduct forms.

Both classes of redox enzymes may proceed in a single step or by two half reactions (akin logically to the PLP-dependent transaminases noted in Section 6.3). In the nicotinamide dependent α -amino acid dehydrogenases, transfer of the C_2 -hydrogen as a hydride ion to C_4 of NAD^+ constitutes the single-step catalytic process. NADH is a two-electron reduced dihydropyridine (see the vitamin-based heterocycles in Chapter 10). The nascent imino acid product is two-electron oxidized (Figure 6.9) relative to the amino acid substrate. As with other imines, the imino acid once released into solution rapidly adds water and the tetrahedral carbinolamine decays rapidly in the forward direction to the observed ketoacid and free NH_3 .

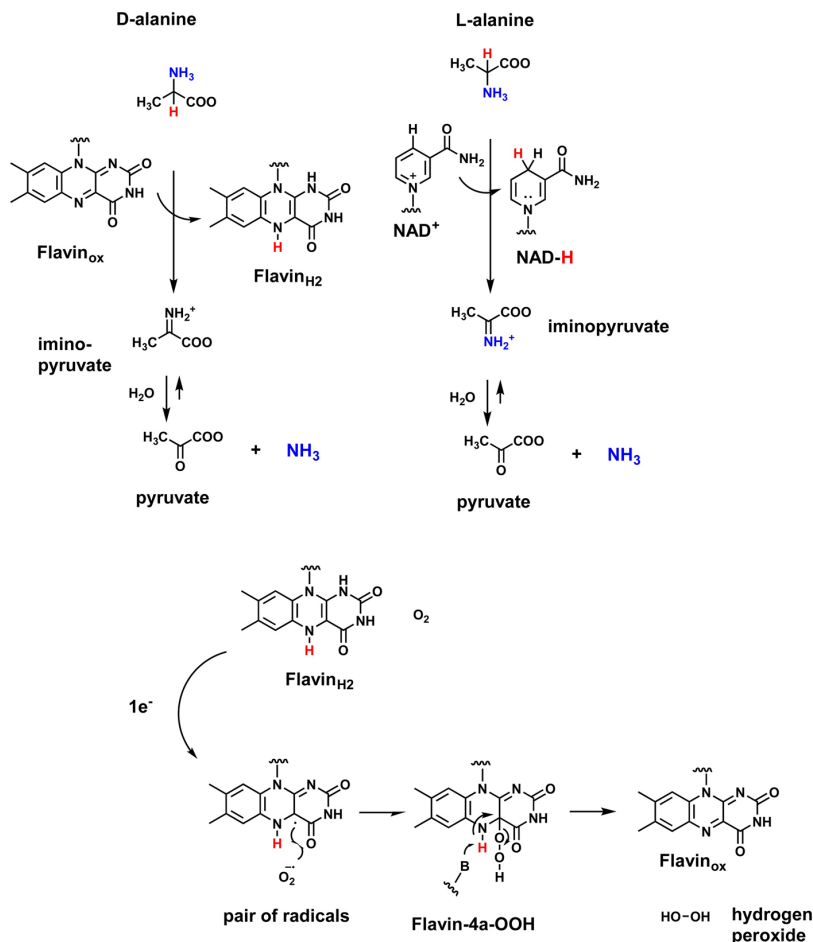


Figure 6.9 Amino acid oxidases generate free ammonia as a secondary product from hydrolysis of initial imino acid products. Two distinct cofactors are used by cells: flavin adenine dinucleotide (FAD) vs. NAD^+ . The NAD^+ is stoichiometrically reduced to NADH. The $FADH_2$ stays tightly bound and is reoxidized by two electrons as O_2 is reduced to H_2O_2 .

For similar enzymes where O_2 is the reducible cosubstrate, H_2O_2 is the typical coproduct along with the imino acid that devolves nonenzymatically to keto acid and NH_3 . However, O_2 as a ground state triplet molecule reacts very slowly with spin-

paired singlet organic scaffolds in all the organic metabolites. (This accounts for our ability to coexist with 21% O₂ in the atmosphere and not undergo instant thermodynamically favored combustion to CO₂ and H₂O.)

Thus, many oxidases contain one equivalent of a tightly bound riboflavin-(vitamin B₂)-based coenzyme **1** (Figure 6.9). Catalysis typically occurs in two half reactions. **21** The substrate oxidative half reaction yields nascent imino acid and two-electron reduced flavin. In the second coenzyme reoxidative half reaction, the dihydroflavin passes its two electrons one at a time to O₂, generating two-electron reduced H₂O₂ as an observed coproduct (Figure 6.9). Kidney L-amino acid dehydrogenases are NAD⁺-dependent dehydrogenases. Snake venom L-amino acid oxidases are flavoenzymes in the toxic venom. Kidney D-amino acid oxidase is also a flavoprotein.

6.4.3 Amidino Group Deaminases: Purines, Pyrimidines and Nucleic Acids

Three additional organic functional groups containing either two or three nitrogen atoms are *ureido* groups (parent compound urea), *amidino* groups (the imine forms of amides) and *guanidino* groups (the imine form of ureas), respectively (Figure 6.10). We defer discussion of ureido groups to the urea cycle in Chapter 7 and of biotin to the chapter on vitamin heterocycles (Chapter 10). Similarly, the guanidinium group of arginine is the immediate progenitor of urea and is examined in Chapter 7 and later chapters. Here, we focus on the enzymatic deaminations of the amidine group, found in the nucleic acid bases that are the building blocks of RNA and DNA.

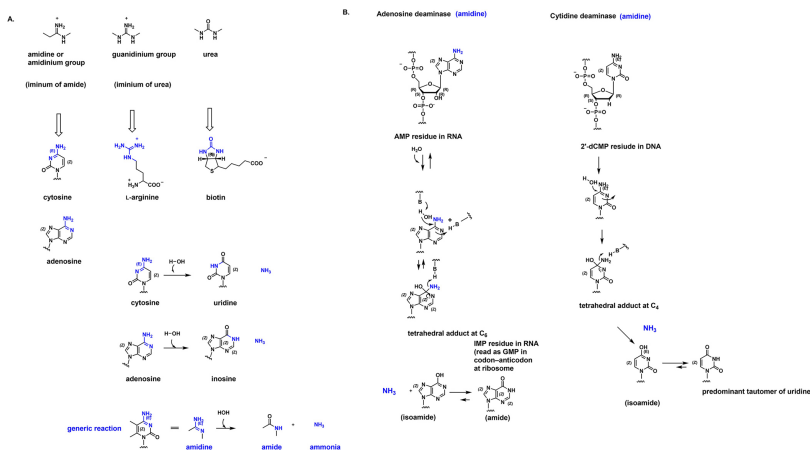


Figure 6.10 Role of amidine, guanidine and ureido groups in organic nitrogen metabolites. (A) Amidines in adenine and cytosine, guanidinium in arginine side chain, urea linkage in biotin (vitamin B₇). (B) Adenosine and cytidine deaminases hydrolyze amidine linkage to amides. A subset of these deaminases operate on adenosine or cytidine residues in RNA and DNA to perform editing functions. The initial isoamides from amidine hydrolyses isomerize to the much more stable amide forms, cause base mispairing.

One pyrimidine, cytosine, and one purine, adenine, contain an amidine linkage. Amidines can be considered as the imine derivatives of amides. In the monocyclic pyrimidine, N₃ and C₄ and the exocyclic amine constitute the amidine linkage, while in the bicyclic purine, the exocyclic amine and the adjacent C=N atoms comprise the

amidine group. Since amidines are the (stable) imine forms of amides, one can imagine they are subject to enzymatic hydrolysis back to the amide and free NH_3 (Figure 6.10). When the cytosine and adenosine moieties are present as dCMP and dAMP residues in DNA the exocyclic amines can function as H bond donors in Watson–Crick H bonds to the antiparallel strand of the DNA double helix. Enzymatic deaminations to uridine or inosine, respectively, convert the pyrimidine and purine instead to H-bond acceptors, altering the H bond polarity and causing miscoding. Some adenosine and cytidine deaminases work on the small molecule nucleosides. Others, with editing functions work on the nucleotides incorporated into RNA or DNA molecules and alter the flow of information transfer.

Adenine deaminase 22 and cytidine deaminase 23 enzymes contain active site Zn^{2+} ions to coordinate both the purine and pyrimidine substrates and to activate the water molecule that functions as nucleophile towards the amidine electrophilic carbon (Figure 6.10B). Water addition to the amidine carbon (C_6 in adenosine, C_4 in cytidine) yields tetrahedral carbinolamine adducts. Breakdown in the forward direction creates the C_6 amide carbonyl of inosine or the C_4 amide carbonyl of uridine, respectively (Figure 6.10B) while releasing free ammonia. Adenosine deaminase deficiency leads to impaired immune system function in humans.

Of particular interest in terms of changing information in either mRNA or DNA are editing versions of the adenine and cytosine bases in the two types of nucleic acids. Two isoforms of human adenosine deaminases 24·25 (ADAR1 and ADAR2) deaminate specific adenine residues to inosine residues in cellular RNAs, 26 both mRNAs and tRNAs. The ketone of the inosine residues is recognized as the ketone of guanosine residues in protein translation at the ribosome and thus the ADARs edit the coding sequence for those RNAs. Aberrant editing is associated or causative in some human diseases such as amyotrophic lateral sclerosis (ALS) and Prader–Willi syndrome.

Comparable cytosine deaminase editing to uridine residues can also occur in both RNAs and also in DNA, 25 changing a C–G pair to a U–A pair on subsequent readout for replication/repair (Figure 6.10B). The human genome encodes six cytidine deaminases 27 in the apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3 (APOBEC3) locus. Those APOBEC3 enzymes deaminate cytidines in single strand DNA regions to “pepper” uridine residues in those stretches of DNA. Among the biological functions are restriction of retrovirus cDNAs and hepadnavirus (e.g. HPV) replications. These newly introduced uridines are also part of the signals for class switch recombination and somatic hypermutation to achieve variability in rearranged immunoglobulin genes. 27 Thus, the biological consequences for these straightforward chemical deaminations are profound. The various editing cytidine deaminases can have different cellular and subcellular locations and different recognition of nucleic acid sequences. 28

6.4.4 Release of NH_3 During Tetrapyrrole Macrocycle Biosynthesis

Most cells have the capacity to make tetrapyrrole macrocycles – from hemoglobin in red blood cells and myoglobin in muscle cells, to chlorophyll in plant chloroplasts and the cobalt containing vitamin B_{12} in microbes. We take up the enzymatic strategy of tetrapyrrole macrocycle biosynthesis in Chapter 11. Here, we note that eight molecules of glycine are fed into the heme biosynthetic pathway. Four of the glycine-derived nitrogens end up as the pyrrole ring nitrogens in the macrocyclic cofactor, while the other four get released as ammonia (Figure 6.11). The responsible deaminase is porphobilinogen deaminase, also known as hydroxylbilane synthase, that forms C–C bonds in pyrrole elongations as each ammonia molecule is released. 29

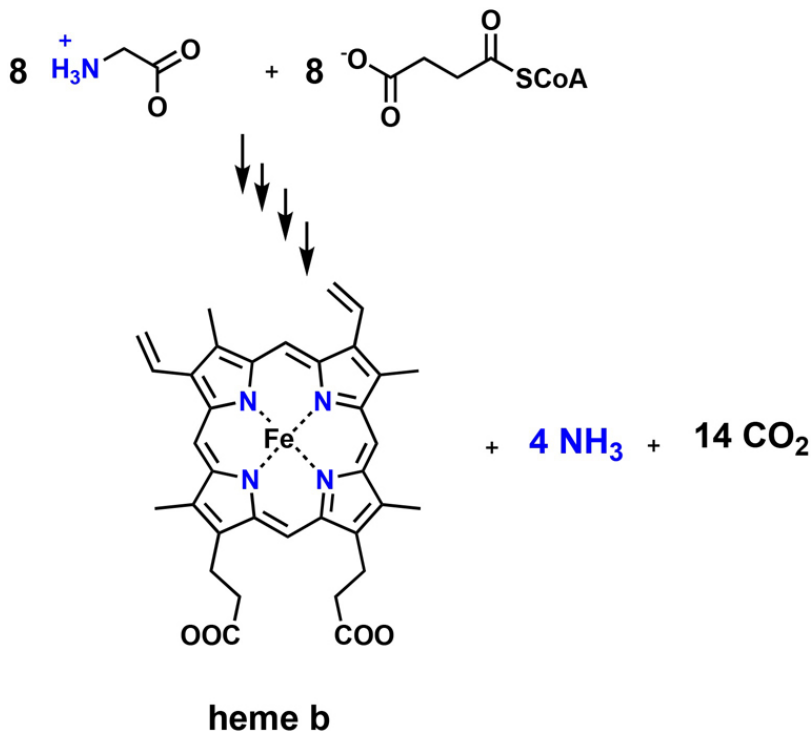


Figure 6.11 The biosynthesis of tetrapyrrole macrocycles such as heme-b in humans uses eight molecules of glycine and of succinyl CoA. Along the way four of the nitrogens are eliminated as NH_3 in the porphyrinogen deaminase or hydroxymethylase synthase step. Also 14 of the 48 starting carbon atoms are jettisoned as 14 CO_2 molecules on the way to the 34-carbon atom framework of heme-b.

The above list of ammonia lyase and deaminative dehydrogenase or oxidase examples reveals the evolution of several different mechanisms to achieve low energy pathways for release of ammonia from amino acids and nucleotide building blocks of RNA and DNA. This topic also serves as a bridge to examination of how cells and organisms manage when toxic levels of NH_3 build up in cells and tissues.

6.5 Overview

NH_3 is the crossover metabolite from the universe of inorganic nitrogen redox metabolism to the realm of dozens to hundreds of organic nitrogen metabolites, both as small molecules and as macromolecules. Glutamate and glutamine are the main entry points. The former is by reduction of the imine of 2-ketoglutarate to the C2-primary amino group of glutamate (Glu). The carboxamido nitrogen of glutamine also arises from ammonia by a typical ATP-dependent amide synthetase reaction. Thus, both the amine and amide nitrogens of glutamine (Gln) come from free ammonia by the two main routes: imine reduction and ATP-dependent amide synthesis.

The Glu–Gln couple then provides itinerant “ NH_3 ” groups to many other scaffolds to enable the metabolic diaspora of amino groups. The amide group of Gln offers up its latent NH_3 moiety by glutaminase action in a range of pathways where the nascent NH_3 moiety in flight ends up in amide, amidine, imine or amine, hemiaminal and aryl

amine linkages. Glutamate mobilizes its amino group by a different logic, utilizing transaminases to shuttle that amino group onto many different 2-keto acid frameworks to populate the cellular pools of both proteinogenic and nonproteinogenic amino acids.

We take up the biosynthesis of nitrogen-containing tetrapyrrolic macrocycles, nitrogen heterocycles in coenzyme forms of vitamins and the biosynthetic logic for pyrimidine and purine heterocycles in three later chapters, building on the dual imine reduction and amide synthesis chemical logic analyzed in this chapter. Before doing so, we continue the theme of how cells deal with free ammonia, liberated *inter alia* by the kinds of deaminases enumerated in this chapter.

References

1. C. T. Walsh and Y. Tang, *Biosynthesis of Natural Products*, Royal Society of Chemistry Press, London UK, 2019.
2. S. H. Liaw and D. Eisenberg Structural model for the reaction mechanism of glutamine synthetase, based on five crystal structures of enzyme-substrate complexes, *Biochemistry*, 1984, **33**, 675-681.
3. C. T. Walsh, *The Chemical Biology of Phosphorus*, Royal Society of Chemistry Press, London UK, 2021.
4. C. T. Walsh, B. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
5. C. Frieden, L-glutamate dehydrogenase, in *The Enzymes VII*, ed. P. D. Boyer, H. Lardy and K. Myrback, Academic Press, 1963, pp. 3-24.
6. M. C. Walker and W. van der Donk The many roles of glutamate in metabolism, *J. Ind. Microbiol. Biotechnol.*, 2015, **43**, 419-430.
7. C. Masclaux-Daubresse, M. Reisdorf-Cren, K. Pageua, M. Lelandais, O. Grandjean, J. Kronenberger, J.-H. Valadier, M. Feraud, T. Joulrt and A. Suzuki Glutamine synthetase-glutamate synthase pathway and glutamate dehydrogenase play distinct roles in the sink-source nitrogen cycle in tobacco, *Plant Physiol.*, 2006, **140**, 444-456.
8. M. A. Vanoni and B. Curti Structure-function studies of glutamate synthases: a class of self-regulated iron-sulfur flavoenzymes essential for nitrogen assimilation, *IUBMB Life*, 2008, **60**, 287-300.
9. F. M. Raushel, J. Thoden and H. M. Holden The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia, *Biochemistry*, 1999, **38**, 7891-7899.
10. F. Massiere and M. Badet-Denisot The mechanism of glutamine-dependent amidotransferases, *Cell Metab. Life Sci.*, 1998, **54**, 205-222.
11. F. M. Raushel, J. Thoden and H. M. Holden Enzymes with molecular tunnels, *Acc. Chem. Res.*, 2003, **36**, 539-548.
12. H. M. Holden, J. Thoden and F. M. Raushel Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product, *Cell. Mol. Life Sci.*, 1999, **56**, 507-522.
13. C. T. Walsh, S. Haynes and B. Ames Aminobenzoates as building blocks for natural product assembly lines, *Nat. Prod. Rep.*, 2011, **29**, 37-59.
14. T. M. Larsen, S. Boehlein, S. M. Schuster, N. G. Richards, J. B. Thoden, H. M. Holden and I. Rayment, *Biochemistry*, 1999, **38**, 16146-16157.
15. B. W. Poland, M. Silva, M. A. Serra, Y. Cho, K. H. Kim, E. M. Harris and R. B. Honzatko Crystal structure of adenylosuccinate synthetase from *Escherichia coli*, *J. Biol. Chem.*, 1993, **268**, 25334-25342.
16. M. Tsai, J. Koo, P. Yip, R. F. Colman, M. L. Segall and P. L. Howell Substrate and product complexes of *E. coli* adenylosuccinate lyase provide new insights into the

- enzymatic mechanism, *J. Mol. Biol.*, 2007, **370**, 541-554.
17. P. K. Mehta and P. Christen The molecular evolution of pyridoxal-phosphate-dependent enzymes, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1999, **74**, 129-184.
18. R. E. Viola The ammonia-lyases: enzymes that use a wide range of approaches to catalyze the same type of reaction, *Crit. Rev. Biochem. Mol. Biol.*, 2019, **54**, 467-483.
19. C. T. Walsh, *Posttranslational Modification of Proteins: Expanding Nature's Inventory*, Roberts & Co, Englewood, Colorado, 2006.
20. B. W. Shen, M. Henning, E. Hohenester, J. N. Jansonius and T. Schirmer Crystal structure of human recombinant ornithine aminotransferase, *J. Mol. Biol.*, 1998, **277**, 81-102.
21. P. Motta, G. Molla, L. Pellegioni and M. Nardinin Structure-function relationships in L-amino acid deaminase, a flavoprotein belonging to a novel class of biotechnologically relevant enzymes, *J. Biol. Chem.*, 2016, **291**, 10457-10475.
22. P. E. Daddona and W. Kelly Human adenosine deaminase. Purification and subunit structure, *J. Biol. Chem.*, 1977, **252**, 110-115.
23. J. Laliberte and R. Mompalao Human cytidine deaminase: purification of enzyme, cloning, and expression of its complementary DNA, *Cancer Res.*, 1994, **54**, 5401-5407.
24. R. A. Mizrahi, K. Phelps, A. Y. Ching and P. A. Beal Nucleoside analog studies indicate differences between RNA-editing adenosine deaminases, *Nucleic Acids Res.*, 2012, **40**, 9825-9835.
25. B. R. Bass RNA editing by adenosine deaminases that act on RNA, *Annu. Rev. Biochem.*, 2002, **70**, 817-846.
26. G. Chen, D. Katrekar and P. Mali RNA-guided adenosine deaminases: advances and challenges for therapeutic RNA editing, *Biochemistry*, 2019, **58**, 1947-1957.
27. R. Suspene, M. M. Aynaud, D. Guetard, M. Henry, G. Eckhoff, A. Machio, P. Pineau, A. Dejean, J.-P. Vartanina and S. Wain-Hobson Somatic hypermutation of human mitochondrial and nuclear DNAs by APOBEC3 cytidine deaminases, a pathway for DNA catabolism, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4858-4863.
28. T. V. Silvas and C. A. Schiffer APOBEC3s: DNA-editing cytidine deaminases, *Protein Sci.*, 2019, **28**, 1552-1566.
29. R. Gill, S. E. Kolstoe, F. Mohammed, A. Al d-Bass, J. E. Moseley, M. Sarwar, J. B. Cooper, S. P. Wood and P. M. Shoolingin-Jordan Structure of human porphobilinogen deaminase at 2.8 angstroms: the molecular basis of acute intermittent porphyria, *Biochem. J.*, 2009, **420**, 17-25.

When Ammonia Becomes Toxic: Ureido and Guanidinium Forms of Organic Nitrogen

7.1 Maintaining a Steady State Nontoxic Level of Ammonia

Organisms from viruses to bacteria to fungi to plants to humans all need nitrogen to survive. Dietary intake of a protein-rich diet leads to digestion of proteins to small peptide fragments and ultimately free amino acids. Those can undergo the variety of deaminase-mediated mechanism leading to release of free ammonia noted in Chapter 6. Because of its basicity and nucleophilicity, high levels of free ammonia can be toxic and are closely regulated by organisms. The main regulatory logic in humans is to excrete ammonia as a nonreactive, non-nucleophilic organic derivative, *urea*.

Ammonia levels in blood can range from 10 to 50 $\mu\text{g dL}^{-1}$. At pH 7.4 in blood, ammonium ions constitute about 99% of the total, with deprotonated NH_3 at about 1%. Thus, free NH_3 is in the range of 10 to 50 $\mu\text{g L}^{-1}$ of blood, corresponding to approximately 0.5 to 3 micromolar ammonia in the presence of 50–250 micromolar NH_4^+ ions. In a total volume of approximately 5 liters of blood, that corresponds to 250 μmoles to 1.25 mmoles . At a molecular weight of 17 g mol^{-1} for NH_3 , that corresponds to approximately 3–17 mg of dissolved ammonium ions plus ammonia in blood. ¹

Because higher levels of ammonia can be toxic, humans act to control ammonia levels in blood and tissues quite carefully, largely by a strategy of detoxification of ammonia to the neutral compound urea. They are termed ureotelic organisms. In contrast, most fish are ammoniotelic, excreting excess ammonia directly into the aqueous medium, where it rapidly gets diluted to nontoxic levels. (Fish tanks have populations of ammonia oxidizing bacteria – Chapter 3.) A third strategy (Figure 7.1) is found in many birds, which excrete solid uric acid salts as a route to rid their bodies of excess nitrogen (uricotelic). The excretion of a solid salt allows for water conservation compared with kidney excretion of aqueous urea solutions.

Problem to be solved



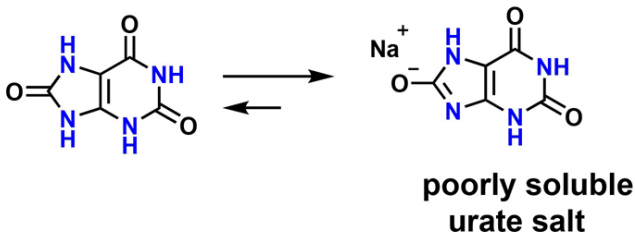
toxicity from nucleophilic reactivity of free NH_3

Three multicellular eukaryotic solutions to eliminate NH_3

1. Fish directly excrete NH_3 into their aqueous medium



2. Birds excrete insoluble salts of uric acid



3. Humans excrete water soluble, unreactive urea at tens of grams per day



Figure 7.1 Different organisms have evolved distinct mechanisms for elimination of toxic high levels of metabolically generated NH_3 .

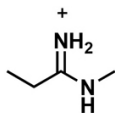
Note that excretion of the poorly soluble urate salt removes four non-nucleophilic nitrogens in one molecule. Some humans with arthritic conditions can have high circulating levels of urate, a metabolic breakdown product from purine building blocks. Under those conditions the poorly soluble urate salts can crystallize in the intracellular spaces of joints, classically around the big toe, causing the painful symptoms of acute gout.

In comparison to urate salts, urea is highly water soluble. The soluble, unreactive urea (carbonyl diamide) is eliminated by the kidney through elaboration of urine. About 5–20 g of ammonia equivalents are excreted daily in the form of urea ($\text{MW} = 50 \text{ g mol}^{-1}$). Urea is highly soluble and neutral so it does not alter blood pH. Both nitrogen atoms in urea are nonnucleophilic, in contrast to ammonia. Blood urea

nitrogen levels are in the range of $2.5\text{--}6.7\text{ mmol L}^{-1} = 125\text{--}400\text{ mg L}^{-1}$ (about 50-fold higher than circulating ammonium ion levels). The average volume of urine elaborated daily by adult humans is in the range of 1.4 liters, giving a value of 12.6 to 32 g of urea in urine and emphasizing the need for daily elimination of ammonia-derived nitrogen atoms.

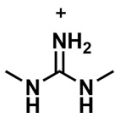
7.2 Amidines, Guanidines and Ureas

Figure 7.2 displays the structure of urea and two related nitrogenous organic part structures, amidines and guanidines. The ureido group is neutral and uncharged at physiological pH values. The lone pair of electrons on each of the two nitrogens are shared with the carbonyl group in isoamide anion part structures, contributing to the unreactivity of the NH_2 groups that have amide like properties.



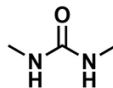
**amidine/
amidinium
group**

**(iminium of
amide)**



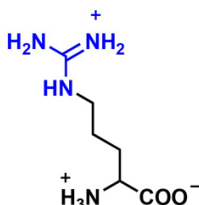
**guanidinium
group**

**(iminium of
urea)**

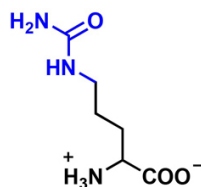


urea

**(simplest
ureido
scaffold)**



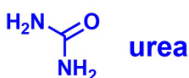
L-arginine



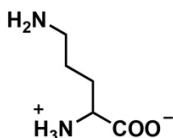
L-citrulline

arginase

HOH



urea



L-ornithine

Figure 7.2 Amidine, urea and guanidinium are three central nitrogen functional groups in the underlying chemical strategy of the urea cycle and removal of ammonia in a nontoxic form.

The guanidino group is the imine derivative of urea. As functional group in the side chain of the proteinogenic amino acid L-arginine, the positively charged guanidinium group has a pK_a value of 12–13. Therefore, at physiological pH values, free in aqueous solutions, the arginine side chain is effectively always positively charged as the guanidinium cation with the positive charge shared between the two terminal nitrogens (Figure 7.2). As a three-nitrogen functional array the guanidinium cationic group is not notably nucleophilic (but Chapter 16 describes several types of posttranslational modifications of arginine residue side chains). It is the iminium ion form of urea. We shall note below that the metabolite citrulline has the ureido part structure (Figure 7.2).

In Chapter 9 on the *de novo* construction of purine and pyrimidine building blocks for RNA and DNA we will return to this metabolic progression of amidine to ureido group to guanidine in the interplay between inosine monophosphate (IMP), xanthosine monophosphate (XMP) and guanosine monophosphate (GMP).

None of the twenty proteinogenic amino acids has a side chain amidine group. We did note in Chapter 6 that the pyrimidine cytosine and the purine adenine both have amidino groups. Those amidines, on the one hand, are key for nucleic acid hydrogen bonding. On the other hand, they are hydrolytically converted to the amides (uracil and inosine) by cytosine and adenosine deaminase actions.

Inspection of the structures of arginine, its biosynthetic precursor ornithine, and the excretable nitrogenous urea, indicate a strategy for elimination of ammonia equivalents as nontoxic urea. Enzymatic hydrolytic cleavage of the guanidinium moiety in the arginine side chain would generate urea and L-ornithine, pinpointing both arginine as the direct metabolic precursor to urea and the key role of the arginase enzyme in the proposed fragmentation (Figure 7.3A). The strategy then devolves to how to build up the guanidinium moiety of arginine by capturing free ammonia. 2

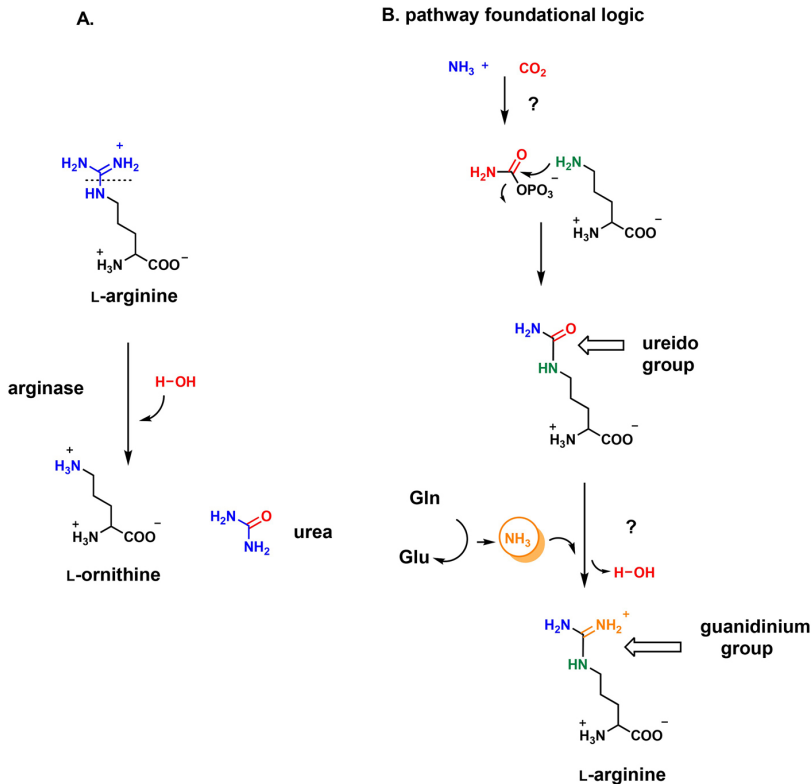


Figure 7.3(A) Urea arises from hydrolysis of the guanidinium side chain of L-arginine by the enzyme arginase. (B) The three nitrogens in the guanidinium group side chain of arginine arise from ornithine, carbamyl phosphate and glutamine: the last two *via* nascent ammonia from glutaminase actions.

This is the function of the urea cycle, starting in human mitochondria. ³ The enzyme carbamyl phosphate synthetase follows the precept similar to conversion of an amine (e.g. NH_3) into a carboxamide linkage. One has to spend an ATP and go through a mixed acyl phosphoric anhydride (carbamyl phosphate, activated for carbamyl transfer to a nucleophile). ⁴ In the urea cycle, the goal is to convert NH_3 into a carbamyl group that can then act as an electrophile to be captured by an organic amine. That step creates a ureido group. The third nitrogen comes from the alpha amino group of aspartate, converting the ureido side chain (in citrulline) to the guanidinium group in arginine (Figure 7.3B).

In the following paragraphs we examine each of the urea cycle enzymes to see how this underlying chemical logic for *de novo* construction of the guanidinium group in arginine is assembled, prior to its hydrolytic cleavage to release urea.

7.3 Carbamate, Carbamyl Phosphate and Carbamyl Transfers

To convert excess ammonia molecules to nontoxic urea for excretion, humans start with the simplest two molecules, free NH_3 and carbon dioxide. From these simple

building blocks, they make an activated form of carbamic acid, the simplest organic acid containing nitrogen. In turn, that activated carbamyl phosphate **4** is a kinetically stable carbamylating agent for transfer onto nucleophilic cosubstrate amines.

It is the arginine biosynthetic pathway in which a ureido group is further aminated to the *guanidinium* moiety in arginine that sets the stage for elimination of urea. Along the way this pathway also illustrates one of the two places in primary metabolism where aspartate is a net amino donor. When urea is finally released, one of the two nonnucleophilic ureido nitrogens came from ammonia, the other from aspartate: two convergent modes of nitrogen excretion.

7.3.1 Carbamic Acid and Carbamyl Phosphate

The formation of carbamic acid can occur by nonenzymatic attack of ammonia on CO₂ as shown in [Figure 7.4](#). In turn, carbamic acid is marginally stable under physiological conditions, reverting back to CO₂ and NH₃. The nonenzymatic condensation is slow at neutral pH where free NH₃ (pK_a = 9.35) is only 1% of the total ammonia–ammonium ion pool. Likewise, CO₂ is the minor component in the CO₂–HCO₃[−] equilibrium, so the bimolecular reaction of nucleophilic NH₃ with electrophilic CO₂ is slow.

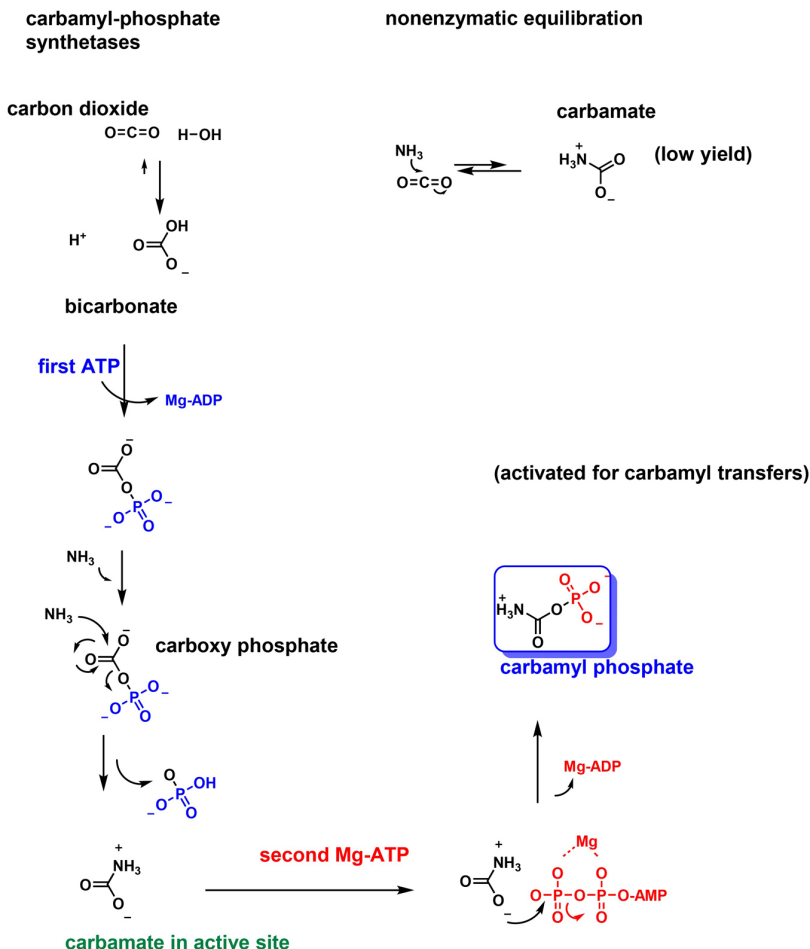


Figure 7.4 Biological strategy: drive accumulation of carbamyl phosphate (a thermodynamically activated carbamylating agent) by progressively cleaving two ATPs for successive phosphoryl transfers in the active site of carbamyl phosphate synthetases producing two sequential acyl phosphates: carboxy phosphate and released carbamyl phosphate.

Instead, animals, and notably humans, in liver tissue, spend an ATP *via* phosphoryl transfer to drive the formation of carbamic acid from NH_3 and HCO_3^- , not CO_2 . Furthermore, the carbamate thus accumulated is retained in the active site of those enzymes and subjected to a second round of ATP cleavage (another phosphoryl transfer) to generate the mixed carbamic-phosphoric anhydride carbamyl phosphate 4 (Figure 7.4). As with all metabolic acyl phosphates carbamyl phosphate is doubly activated: 5 for carbamyl transfer to attacking nucleophiles or for phosphoryl transfer to a different set of nucleophiles (e.g. ADP in the back reaction).

Humans make two isoenzyme forms of these carbamyl phosphate synthetases. 6 One is in the cytoplasm of hepatic and other cells and catalyzes the first step in pyrimidine ring assembly (to be discussed in Chapter 9). The second is in hepatocyte mitochondria, is activated allosterically by *N*-acetylglutamate, 7 and is the first enzyme in a six-enzyme pathway termed the urea cycle, first deciphered by Hams Krebs and Kurt Henseleit in 1932 2:3 (Figure 7.5).

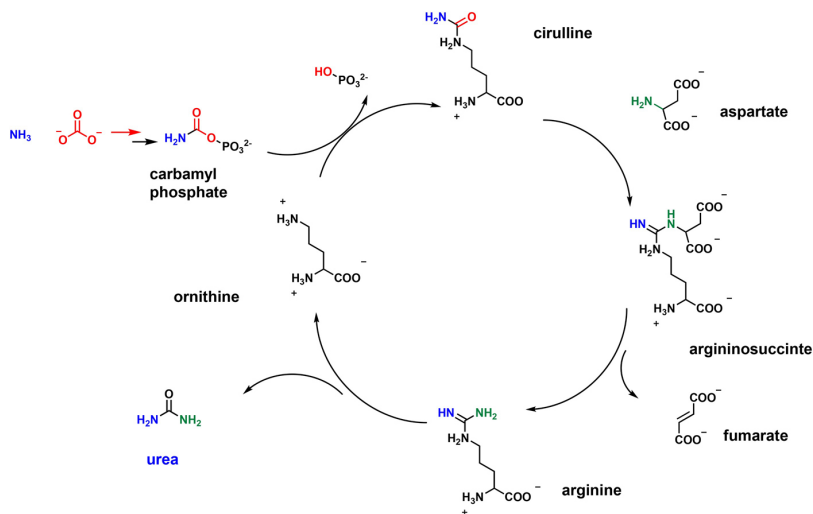


Figure 7.5 The path of nitrogen through the six enzymes of the urea cycle.

Both carbamyl phosphate synthetases use the same catalytic strategy and mechanism. 6 The predominant one-carbon anion bicarbonate is the nucleophilic substrate used to attack Mg-ATP at the electrophilic P_γ phosphorus atom of the ATP side chain. The result is formation of Mg-ADP and the mixed anhydride between carbonic acid and inorganic phosphoric acid, an intermediate termed *carbonyl phosphate* or *carboxy phosphate* (Figure 7.4). This is the simplest of all the acyl phosphate metabolites, thermodynamically activated at both the electrophilic carbon and electrophilic phosphorus atoms. Carboxy phosphate is not stable kinetically and is held tightly in the enzyme active site.

Its fate is to be attacked on carbon by the free base form of ammonia, the third substrate in this first reaction sequence. That forms the N–C bond of carbamic acid

while releasing inorganic phosphate. The first ATP (colored blue in [Figure 7.4](#)) has done its job, providing the driving force for accumulating carbamic acid in the carbamyl phosphate synthetase active site. The preferred source of the nascent ammonia is glutamine, by action of the glutaminase domain of carbamyl phosphate synthetases, thus avoiding high concentrations of soluble, diffusible ammonia.

At this point, with carbamate generated in the active site, the synthetase binds a second molecule of Mg-ATP and again cleaves it to ADP and Pi to drive the equilibrium in the second half reaction in the forward direction. The carboxylate anion of the *in situ* carbamate is the oxygen nucleophile attacking P_γ of the second Mg-ATP (colored red in [Figure 7.4](#)). This generates a second acyl phosphate, in this case carbamyl phosphate. While kinetically labile, (it has half times of minutes not seconds), this thermodynamically activated metabolite is released from the enzyme and lasts long enough *in cellulo* to diffuse to cosubstrates and the next sets of enzymes to serve as the premier cellular carbamylation reagent.

7.3.2 Aspartate Transcarbamylase and Ornithine Transcarbamylase

Two enzymes in distinct primary metabolic pathways utilize carbamyl phosphate as an acylating (carbamylating) agent towards nucleophilic amines of amino acid substrates. Aspartate transcarbamylase uses the amino group of aspartate in cytoplasmic assembly of nucleic acid pyrimidine building blocks. [8](#) Ornithine transcarbamylase [9](#) instead reroutes nitrogen flux down the pathway to arginine and ultimately urea for excretion. [10](#) (In older literature the alternative spelling of *carbamoyl* for *carbamoyl phosphate* and *transcarbamoylases* is often used.)

[Figure 7.6](#) depicts the action of the two transcarbamylases. We will return to the aspartate transcarbamylase in Chapter 9, delineating origins of the nitrogens in the pyrimidine and purine heterocyclic bases of RNA and DNA. Here, we note that when the aspartate amino group acts as a nucleophile towards the activated acyl group carbonyl of carbamyl phosphate, it creates a *ureido* group (–NH–CO–NH–) in the product *N*-carbamyl aspartate.

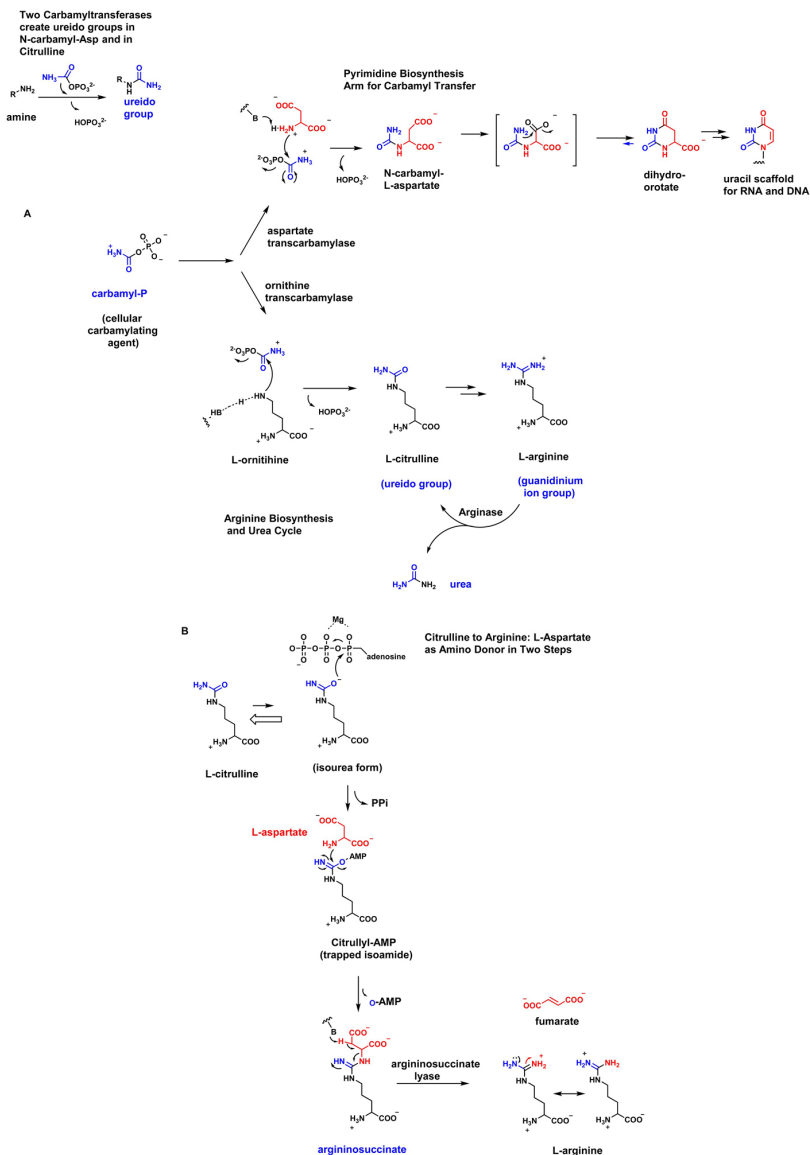


Figure 7.6 Enzymatic amine carbamylations generate urea groups in products. (A) Carbamyl phosphate as branchpoint metabolite to make citrulline on the way to arginine or N -carbamyl-L-aspartate on the way to pyrimidine scaffolds. (B) Citrulline to arginine.

Analogously, the side chain C_5 amine of L-ornithine is directed to attack the carbamyl moiety in the active site of ornithine transcarbamylase. The amino acid product known as citrulline also has a *newly installed ureido group*. These two metabolites are the first ones we have come upon in which amino nitrogens have been converted to deactivated ureido group nitrogens (a carbonyl diamide).

By the way, cellular ornithine is made from glutamate. ¹¹ The C_5 carboxylate of glutamate has to be reduced by four electrons to get to the CH_2NH_2 group of ornithine

at C₅; this occurs in two separate hydride reduction steps from NADH. The first hydride transfer to C₅ of glutamate yields the aldehyde (glutamic-5-semialdehyde). Iminium formation between the aldehyde and NH₃ yields the semialdehyde cationic iminium ion. Use of the second NADH to reduce the imine creates L-ornithine and re-emphasizes the wide utility of reductive aminations, seen also in Chapter 6 with glutamate dehydrogenase. Both that enzyme and this ornithine dehydrogenase-synthase practice reductive aminations.

The next pair of enzymes act in tandem to introduce a third nitrogen atom at the carbonyl of the ureido group of citrulline. When the dust settles that ureido group has been converted to a guanidino group. The pK_a of the guanidine in the arginine side chain is approximately 13 so under physiological conditions the neutral ureido moiety of citrulline has been transformed into the cationic guanidinium side chain of arginine.

7.3.3 Citrulline to Arginine: Ureido to Guanidinium Group via Addition–Elimination of Aspartate C₂-Amino Group

The first of a pair of tandem enzymes converting the ureido group of citrulline to the guanidinium group of arginine has the task of reversing the polarity of the ureido carbonyl oxygen from an electrophile to a nucleophile. The argininosuccinate synthetase enzyme **12** uses a strategy of converting the electrophilic ureido carbonyl of citrulline into the (very) minor isoamide form in the enzyme active site such that the oxyanion is now a nucleophile (**Figure 7.7**). The cosubstrate that gets attacked is again Mg–ATP but now the regioselectivity is controlled such that the isoamide anion attacks the electrophilic P_α rather than P_γ, an altered regioselectivity from carbamyl phosphate synthetase. This constitutes an adenylyl (–AMP) group transfer to the citrulline isoamide oxygen, producing citrullinyl–AMP and releasing inorganic pyrophosphate (**Figure 7.7**).

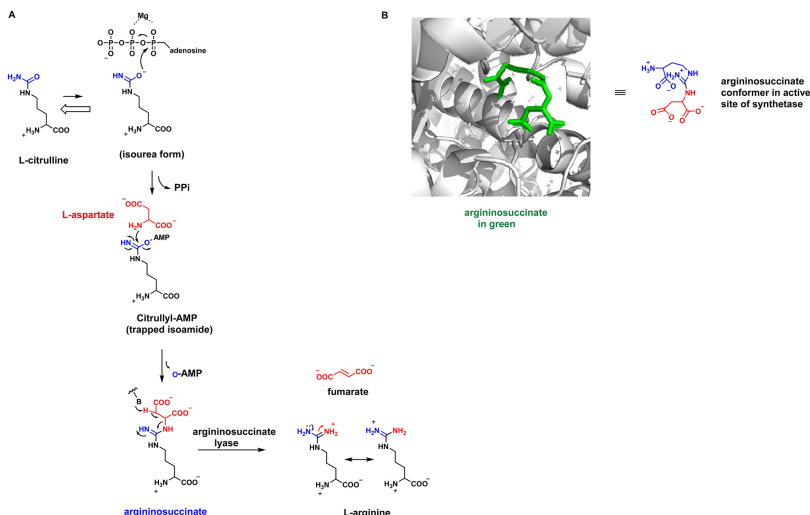


Figure 7.7 (A) The net amination of citrulline to arginine is a two-enzyme process using aspartate as an amino donor by an addition–elimination strategy via the intermediate argininosuccinate. (B) A close-up view of the bound argininosuccinate conformer in the synthetase active site, from PDB 1J20.

The citrullinyl–AMP is held tightly at the active site and in position to be attacked by

cosubstrate aspartate *via* its deprotonated nucleophilic amino group. The oxygen that had initially been the carbonyl oxygen of the ureido moiety in citrulline now departs as part of the AMP product (with a lower energy barrier than if it were to be ejected as an OH^-). The nitrogen from aspartate is now in a bridging position between the citrulline framework and the original four-carbon aspartyl scaffold. This is argininosuccinate. The orientation of the argininosuccinate product before release from the argininosuccinate synthetase active site is shown in [Figure 7.7B](#) (redrawn from pdb file 1J20).

The second of the pair of tandem enzymes is argininosuccinate lyase. [13](#) Its function is to cleave the $\text{C}_2\text{--N}$ bond in the aspartate part of this complex scaffold ([Figure 7.7](#)). The lyase acts by hydrogen abstraction of one of the two prochiral methylene hydrogens at C_3 of the aspartyl moiety. Fumarate (a citrate cycle primary metabolite) is generated as one product. The second product is the guanidinium group in the arginine product.

This strategy to use aspartate as an amino donor can be contrasted with use of glutamine as an amino donor. We noted above that some 16 enzymes have glutaminase activity [13](#) to release nascent ammonia from the carboxamide of glutamine as a relatively straightforward hydrolytic drop of nucleophilic ammonia into a target enzyme's active site.

For aspartate the net transfer of its amino group by the two-step addition–elimination process is more baroque. [Figure 7.8](#) notes that the series of three enzymes, ornithine transcarbamylase, argininosuccinate synthase and argininosuccinate lyase, represent a two-stage conversion of amine to guanidinium group. The carbamylation of an amine generates a ureido functionality. The more difficult step is insertion of the third nitrogen, requiring the reversal of polarity of the ureido carbonyl oxygen and its ultimate extrusion after aspartate has added. That reversal is achieved by trapping the rare isoamide anion as the isoamide–AMP adduct before aspartate addition. Finally, the elimination of fumarate harks back to the aspartate ammonia lyase–deaminase reaction of [Figure 6.8](#). In that case fumarate and free NH_3 were the products. In this situation the NH_2 left behind is covalently tethered in the newly formed guanidinium group.

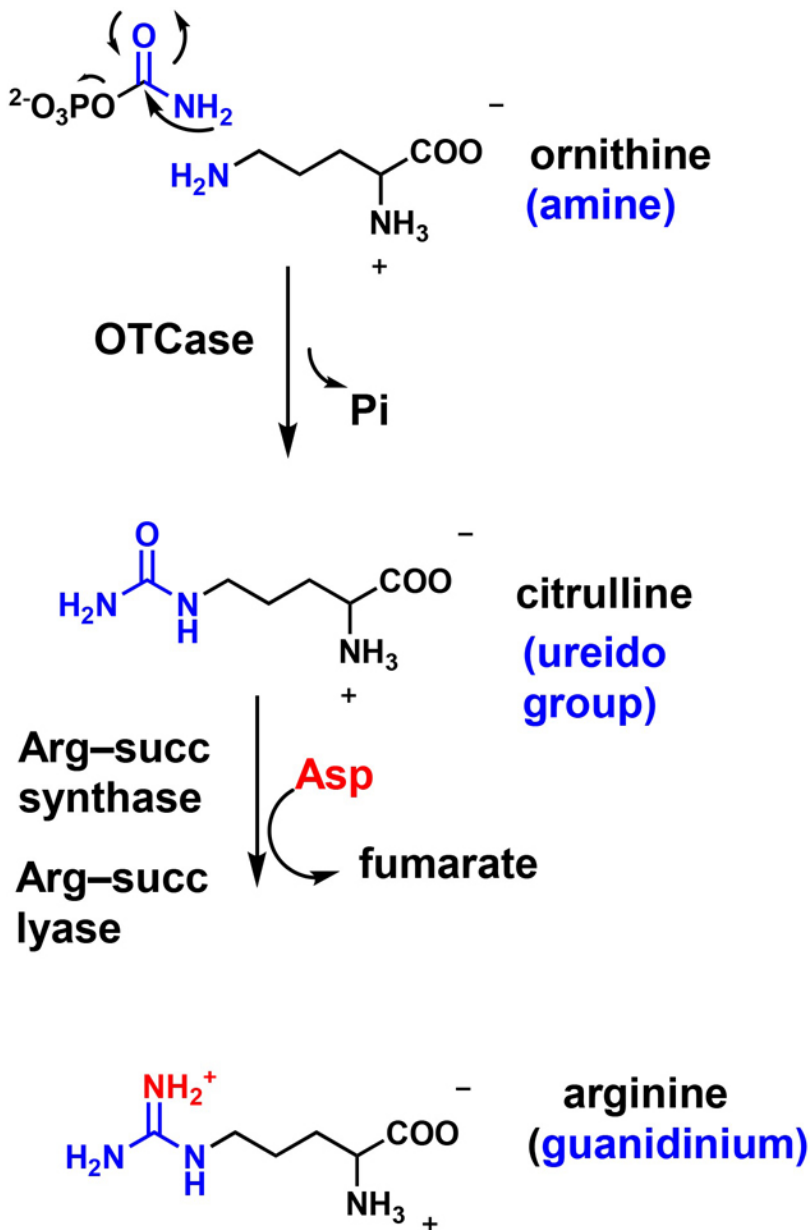
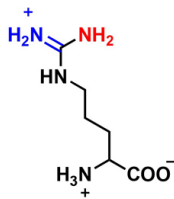


Figure 7.8 Amine carbamylation to ureido group; amination to amidine: three successive enzymes in conversion of ornithine to arginine.

7.3.4 Arginase Completes the Urea Cycle

The last enzyme in the pathway that jettisons two ammonia equivalents as unreactive, highly soluble urea is arginase ¹⁴. Mammalian cells have two arginase isoenzymes, one serves to balance the cellular levels of arginine vs. ornithine. ¹⁵ The second isozyme is localized in mitochondria where it is involved in excretory urea formation.

Both arginases have active site Mn^{2+} ions that act as a binuclear cluster to activate the water molecule cosubstrate to act as an OH^- equivalent on the guanidinium side chain of arginine (Figure 7.9). The central carbon of the guanidinium group is the electrophilic center attacked to yield the indicated tetrahedral adduct. Its breakdown involves a C–N bond cleavage, releasing ornithine and urea. The five-carbon ornithine, with C_2 and C_5 amine groups, is the substrate for ornithine transcarbamylase, so its formation by arginase closes the metabolic urea cycle.



arginase to urea:
active site binuclear
 Mn^{2+} cluster

L-arginine

(one N from carbamyl phosphate
one N from aspartate)

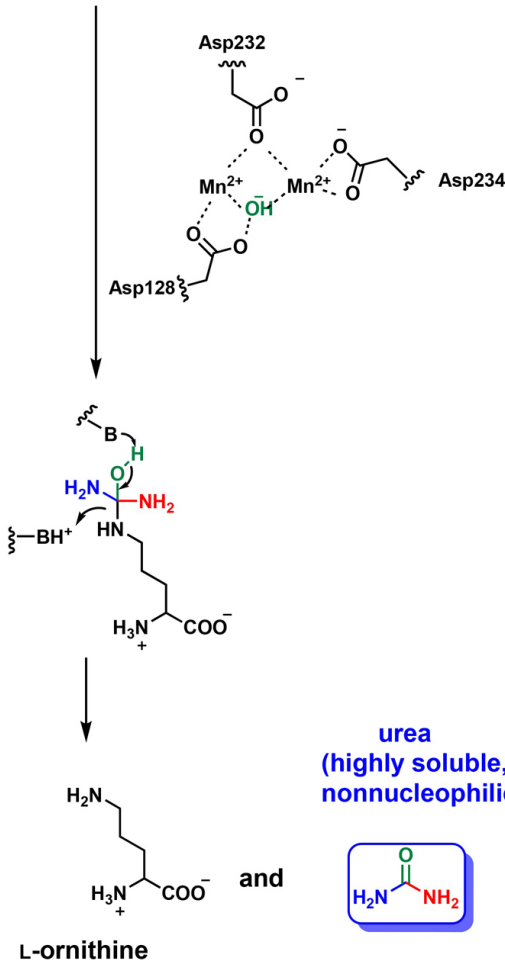


Figure 7.9 Arginase is a bi-zinc-containing guanidium group hydrolase, the last enzyme in the urea cycle, creating free urea and ornithine.

As a measure of flux to eliminate NH_3 from the body, up to 10 g of urea are

generated by humans each day, released into the blood, filtered in the kidney and excreted in urine daily. From an underlying chemical perspective, the use of the acyl phosphate carbamyl phosphate as a carbamylating agent is one precept that emerges. The second is that carbamylation of amines generates ureas. The third precept is that the urea to guanidinium conversion requires ATP cleavage to trap an isoamide-AMP as the target for adding in the third nitrogen as a nucleophilic amine. Fourth, the pathway logic reveals that aspartate can act as an amino donor and thus joins glutamine as the two common amino donors to electrophilic cosubstrates, albeit by very different mechanisms.

7.3.5 Amidino Transfer from Arginine to Creatine

Once the amidino group is created in the arginine framework it can be mobilized for transfer to other acceptors. Most notable in human metabolism is the enzyme-catalyzed formation of creatine from glycine [16](#) ([Figure 7.10](#)). The amine group of the two-carbon proteinogenic amino acid glycine is the nucleophile towards the amidino group while it is covalently tethered to an active site cysteine thiol group.

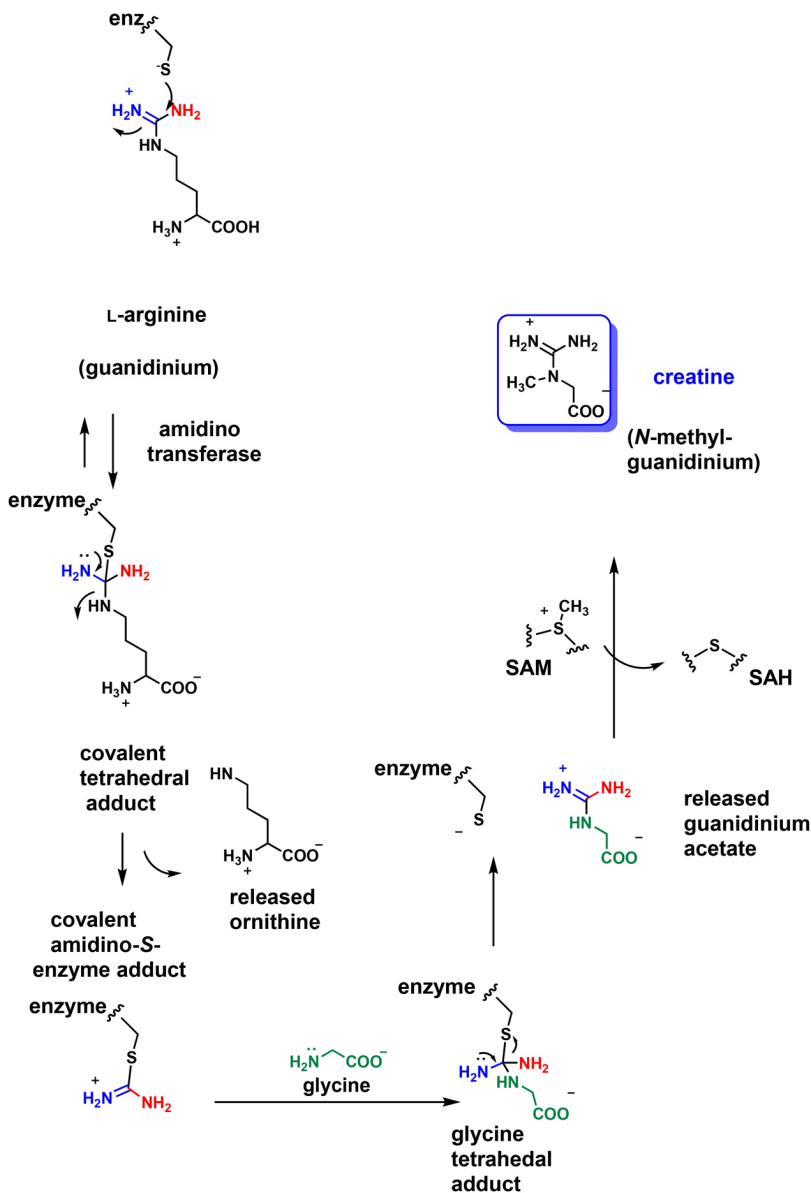


Figure 7.10 Amidino group transfer: arginine to guanidinoacetate on the way to creatine via a covalent amidino-S-enzyme intermediate during the amidinotransferase catalytic cycle. Amidino transfer to an amine generates a guanidine group.

Mobilization of the amidino group from arginine nonhydrolytically is initiated by attack of an active site cysteine thiolate side chain as shown to release ornithine and produce an amidino-S-enzyme covalent adduct; that is then captured by glycine. The resulting guanidinoacetate metabolite is an *N*-methylation step away from forming creatine.

Levels of creatine are high in skeletal muscle where it undergoes reversible *N*-

phosphorylation on the terminal amine of the guanidino group. The result is *N*-phosphocreatine a so-called *phosphagen*, that acts as a thermodynamically activated reservoir to form ATP from ADP (Figure 7.10) and provide ATP quickly e.g. for sustained muscle contraction. 16 We will return to creatine kinase in subsequent discussion of *N*-phosphorylation reactions. Creatine thus introduces amidino group transfer enzymology and *N*-methylation to the secondary amine.

7.4 Overview

This chapter has introduced metabolic progressions from amines to ureas to guanidines. Analogous transformations convert amines to amidines. Carbamyl phosphate, formed at the expense of two ATPs cleaved to drive the forward equilibria, is the cellular carbamylating reagent. Amines that undergo carbamylations yield substituted ureas as products. The chemical progression through these common organic nitrogen groups, with no change in nitrogen redox state, are mainstays of biological logic in other pathways, including for nucleic acid building blocks taken up in Chapter 9.

References

1. I. D. Weiner, W. Mitch and J. M. Sands Urea and ammonia metabolism and the control of renal nitrogen excretion, *Clin. J. Am. Soc. Nephrol.*, 2015, **10**, 1444-1458.
2. D. A. Atkinson Functional roles of urea in vertebrates, *Physiol. Zool.*, 1991, **65**, 243-267.
3. *Nitrogen Metabolism and the Urea Cycle*, info@themedicalbiochemistrypage.org.
4. H. M. Holden, J. Thoden and F. M. Raushel Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product, *Cell. Mol. Life Sci.*, 1999, **56**, 507-522.
5. C. T. Walsh, B. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
6. F. M. Raushel, J. Thoden and H. M. Holden The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia, *Biochemistry*, 1999, **38**, 7891-7899.
7. L. Caldovic, N. AhMew, D. Shi, H. Morizono, M. Yudkoff and M. Tuchman *N*-Acetylglutamate synthase: structure, function and defects, *Mol. Genet. Metab.*, 2010, **100**, S13-S19.
8. J. E. Gouaux and W. Lipscomb Three dimensional structure of carbamoyl phosphate and succinate bound to aspartate transcarbamoylase, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 4205-4208.
9. C. J. Lusty, R. Jilka and E. Nietsch Ornithine transcarbamylase from rat liver. Kinetic, physical and chemical properties, *J. Biol. Chem.*, 1979, **254**, 10030-10036.
10. R. Cunin, N. Glansdorff, A. Pierard and V. Stalon Biosynthesis and metabolism of arginine in bacteria, *Microbiol. Rev.*, 1986, **50**, 314-352.
11. S. M. Morris Enzymes of arginine metabolism, *J. Nutr.*, 2004, **134**, 2743S-2747S.
12. C. Ghose and F. M. Raushel Determination of the mechanism of argininosuccinate synthetase by static and dynamic quench experiments, *Biochemistry*, 1985, **24**, 5894-5898.
13. L. Sampaleanu, B. Yu and P. I. Howell Mutational analysis of duck crystallin and the structure of an inactive mutant with bound substrate provide insight into the

- enzymatic mechanism of argininosuccinate lyase, *J. Biol. Chem.*, 2002, **277**, 4166-4175.
14. D. E. Ash Structure and function of arginases, *J. Nutr.*, 2004, **134**, 2760S-2764S.
15. R. B. Caldwell, H. Toque, S. P. Narayanan and R. W. Caldwell Arginase: an old enzyme with new tricks, *Trends Pharmacol. Sci.*, 2016, **36**, 395-405.
16. J. B. Walker Creatine biosynthesis, regulation and function, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1979, **50**, 177-242.

Metabolic Capture of Amines, Imines and Enamines

8.1 Amines, Imines and Enamines

Chapter 6 and 7 together examined the scope of metabolism of nucleophilic NH_3 in protonic equilibrium, with NH_4^+ cations that are not directly reactive. Chapter 6 examined fixation of NH_3 into organic scaffolds by two major routes: imine formation with the 2-keto acid oxoglutarate to generate glutamate or ATP-dependent amide formation by glutamine synthetase.

To avoid high levels of indiscriminately reactive NH_3 , organisms, including humans, have evolved two different strategies. One is to use the carboxamide group of glutamine as a stable carrier of latent ammonia to be delivered by controlled hydrolysis for use within the interior of targeted amination enzymes. ¹ A complementary strategy, summarized in Chapter 7, is to keep NH_3 levels low by capturing ammonia in the form of the amidine group in citrulline and then the guanidinium group in arginine before arginase-mediated cleavage to yield highly water-soluble urea with two unreactive nitrogens for kidney excretion.

In this chapter we make further examination of some common metabolic fates of organic amines (e.g. in both the proteinogenic and nonproteinogenic amino acid pools), of imines and of enamines (Figure 8.1). The organic metabolite amines, when deprotonated from the ammonium cation forms that predominate at physiological pH values, behave as nucleophiles at nitrogen. The imines, carbonyl adducts with amines, especially in the iminium forms protonated on nitrogen, behave as carbon electrophiles. For imines that arise from aldehydes or ketones where the beta carbon has at least one C-H bond, the enamine tautomers can behave as beta carbanion equivalents. ² Thus, the trio of amines, imines and enamines in metabolite scaffolds account for much of organic nitrogen metabolism. We start with amine reactions, then proceed to a second look at imines, as precursors to enamine chemical biology.

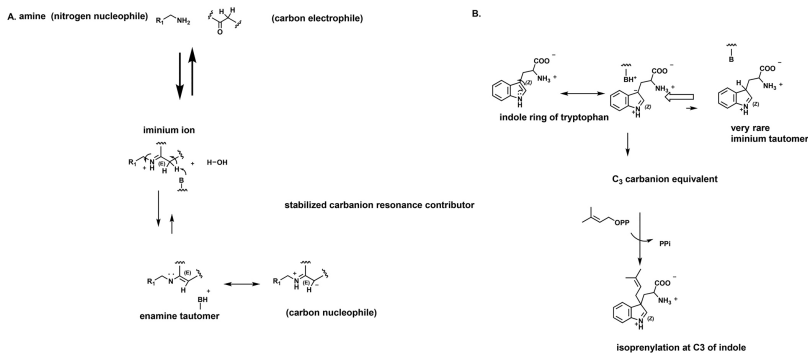


Figure 8.1 (A) Imine–enamine tautomerization in aldimines or ketimines with β -hydrogen substituents yields stabilized carbanions as carbon nucleophiles. (B) The indole ring in the side chain of tryptophan has a built-in enamine part structure. Shown is prenyl transfer to a C_3 carbanion equivalent on the indole ring.

8.1.1 Capture of Substrate Amino Groups in Electrophilic Group Transfer Reactions

One could classify amine metabolite reactions based on analysis of particular metabolic pathways. We take that approach specifically around the logic of purine and pyrimidine ring constructions in Chapter 9. In this chapter we instead note that major categories of enzymes transfer electrophilic fragments of substrates to nitrogen and oxygen nucleophiles in cosubstrates. [3](#)

These categories include: (1) the many hundreds of *phosphoryl* and *nucleotidyl* transfers from ATP and GTP; (2) all the *acyl group* transfers to both low molecular weight nucleophilic metabolites and all proteases; (3) *glycosyl* transfers in monosaccharide, oligosaccharide and polysaccharide assemblies and also in *N*- and *O*-glycoprotein modifications; (4) *alkylation* reactions that include many dozens of one carbon transfers from *S*-adenosylmethionine, and thousands of prenyl group transfers in isoprenoid metabolism; (5) *sulfuryl* group transfers; (6) a subset of *oxygenations* where the substrate undergoing oxygenation is nucleophilic.

[Figure 8.2](#) depicts the generic reactions for these six major electrophilic group transfers that encompass hundreds of metabolic transformations in cells. The underlying chemical logic for all these reaction types is that cells have a set of common reagents that offer *thermodynamically activated but kinetically stable electrophilic groups* to cosubstrate nucleophiles. [3](#) These key reagents include acyl coenzyme A (CoA) thioesters, the sulfonium cation in *S*-adenosylmethionine, the side chain phosphoric anhydrides in ATP, the corresponding mixed phosphoric–sulfuric anhydride linkage in AMP–sulfate, the activated glucosyl moiety in UDP glucose and even molecular oxygen (O_2). These six reaction types involve chemically different functional groups, but all are united in their ability to be captured by cosubstrate nucleophiles. Prominent among those nucleophiles are the many cases where RX = a nucleophilic nitrogen. Those reactions are the topic of this chapter.

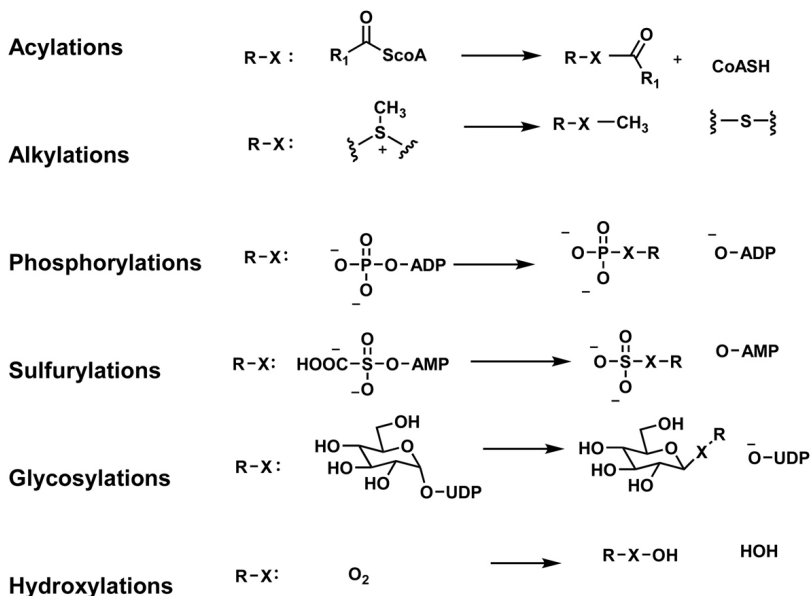


Figure 8.2 Six major categories of electrophilic group transfers to nucleophilic partners (RX).

8.1.2 Amine Acylations

Activated acyl groups are typically carried as thioesters with the acyl group attached to the thiol end of coenzyme A when the acyl groups are short chain. The longer fatty acyl chains such as C₁₄ myristoyl or C₁₆ palmitoyl are assembled on the thiol group of the covalently tethered phosphopantetheinyl arm [4](#) of fatty acid synthase and then offloaded as myristoyl CoA and palmitoyl CoA. The most abundant short chain acyl CoA thioesters in cells are acetyl CoA, malonyl CoA, succinyl CoA and 3-OH-butyryl CoA. Acetyl CoA is the predominant substrate for acyl transfers to cosubstrate amines.

The products from *N*-acylations are amide linkages. We have seen in Chapter 6 that amide formation requires energy input to convert unreactive carboxylate anions to acyl-phosphates or acyl-AMPs. Those species are intermediates in the assembly of acyl CoA thioesters. The acyl CoAs maintain the thermodynamic activation of the acyl-phosphates but offer additional kinetic stability to serve as diffusible molecules throughout cells. [3](#)

The example of glutamate conversion to *N*-acetylglutamate is shown in [Figure 8.3](#). *N*-acetylglutamate [5](#) is an activating ligand for the mitochondrial carbamyl phosphate synthase (Chapter 7) that starts the flux of ammonia into urea for elimination. We will take up lysine side chain *N*-acetylation in Section IV, Chapter 16, on nitrogen proteomics.

Metabolic Reactions of Amino Groups
with Electrophilic Molecules

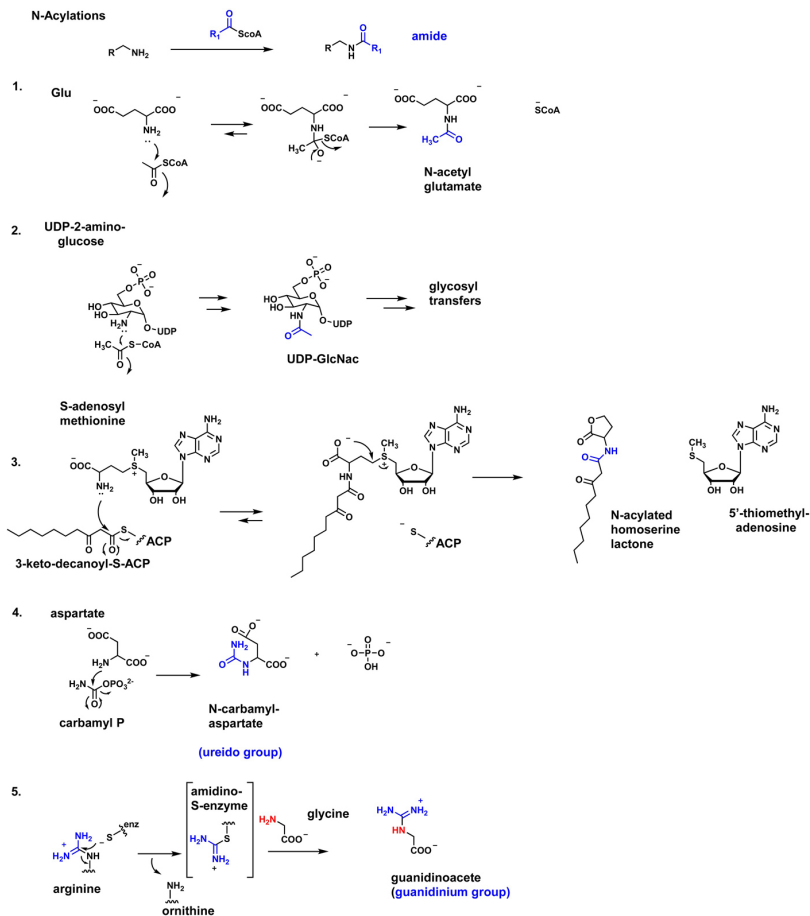


Figure 8.3 Five examples of reactions of amines as nucleophiles with electrophilic substrates: acetyl CoA, 3-keto-decanoyl CoA, S-adenosylmethionine, carbamyl phosphate, arginine amidino group.

Three other metabolic amine acylations are also summarized in [Figure 8.3](#). One is the *N*-acetylation of UDP-glucosamine to *N*-acetylate the aminoglucose moiety to dampen down its nucleophilic reactivity. Acetyl CoA is again the acetyl group donor. UDP-GlcNAc is subsequently the donor for all protein *N*-glycosylation posttranslational modifications of hundreds to thousands of human secreted proteins (Chapter 16).

A third example involves bacterial metabolism in the biosynthesis of a quorum signaling molecule 3-keto-decanoyl-homoserine lactone. [6](#) The C₁₀ fatty acid moiety is an intermediate in fatty acid biosynthesis, tethered in thioester linkage to the acyl carrier protein (ACP) subunit of the bacterial fatty acid synthase. The amino group of the methionyl side chain of S-adenosylmethionine attacks the activated C₁₀ acyl group to form the product amide linkage. The immediate product is the *N*-acylated-S-adenosylmethionine (SAM) but that reacts further before release from the active site. The carboxylate anion acts as internal nucleophile at C₄ of the aminobutyryl chain of this *N*-acylated SAM to release thiomethyladenosine and form the cyclic homoserine lactone. The 3-keto-decanoyl-homoserine lactone is the released product and is a

metabolite that signals high density in bacterial population-quorum sensing.

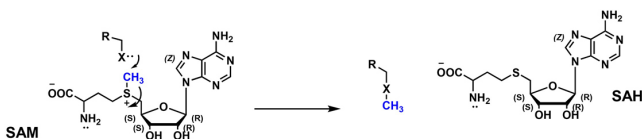
The fourth case of *N*-acylation in Figure 8.3 is the carbamylation of the aspartate amino group to form *N*-carbamyl aspartate 7 (Figure 7.6). This example serves as a reminder that the carbamyl group in this acyl phosphate is the key acylation reagent for both the urea metabolic cycle and also in initiation of pyrimidine biosynthesis.

Figure 8.3 includes one more example of an amino acid, glycine, acting as nitrogen nucleophile but the group undergoing transfer as an electrophilic fragment is not a simple acyl group. Instead, it is the iminium derivative of an amide, the *amidino* group. Net amidino transfer creates the guanidinium group in guanidinoacetate, an *N*-methylation away from creatine 8 (see Chapter 7 Section 7.3.5). This also serves as a segue to the next type of amine capture by electrophiles, *N*-alkylations.

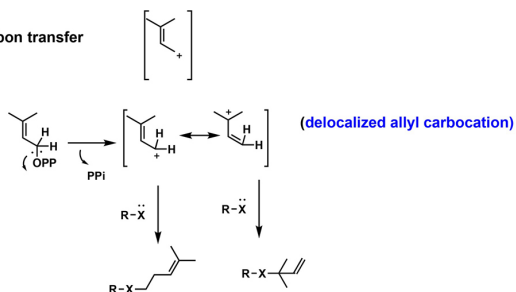
8.1.3 Metabolic Amine Alkylations

There are two distinct types of metabolites that circulate in cells and act as donors of electrophilic alkyl groups, either as a one carbon $[\text{CH}_3]^+$ equivalent or as a donor of branched-chain five-carbon isoprene cations (or multiples thereof, e.g. C_{10} , C_{15} allyl cations) (Figure 8.4). The one-carbon donor is *S*-adenosylmethionine. While the permanent positive charge on the sulfonium sulfur activates all three carbon substituents as electrophilic (cationic) equivalents, methyl group transfer as $[\text{CH}_3]^+$ to O, N, S, or C nucleophiles is by far the most prevalent. Stereochemical studies with C-HDT chiral methyl groups indicate that SAM donates its methyl group in a direct $\text{S}_{\text{N}}2$ type of displacement with inversion of stereochemistry at the one-carbon unit in flight.

1. One-carbon transfer $[\text{CH}_3^+]$



2A. Five-carbon transfer



2B. Ten-carbon transfer

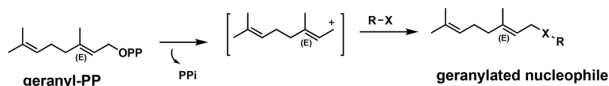
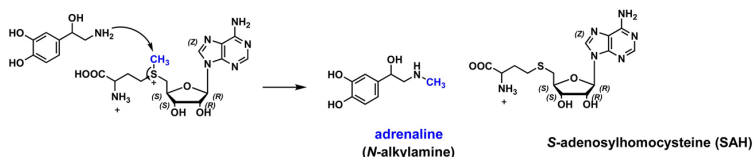


Figure 8.4 Types of cellular alkyl donors: One carbon from *S*-adenosylmethionine, Five carbons and multiples thereof from prenyl diphosphates.

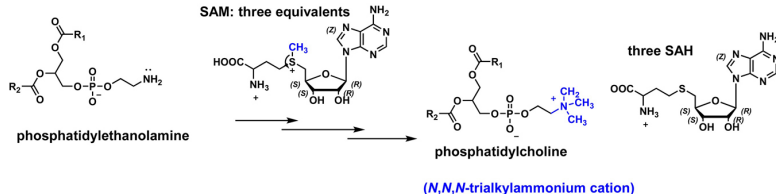
The prenyl group of dimethylallyl diphosphate (DMAPP) is the building block for more than 50 000 known isoprenoid natural products, including all steroids. 3.4 The chemical feature of DMAPP that engenders such biological utility is the fact that the C_{2,3}-olefin sets up a delocalized, stabilized allylic carbocation once C₁–OPP carbon–oxygen bond cleavage occurs. That cleavage is an early event in prenyl transferase transition states, creating an accessible C₁ carbenium ion. Also of note is the delocalization that adds stability, which also means there is a cationic resonance contributor at C₃, accounting for observed “reverse” prenyl transfer regiochemistry when the five carbon cation is captured at C₃ rather than C₁.

Six examples of enzymatic *N*-alkylation reactions are collated in Figures 8.5 and 8.6. The first is a classic *N*-methylation of a primary amine by [CH₃⁺] group transfer from *S*-adenosylmethionine. In this case it depicts the final step in biosynthesis of the adrenal hormone and neurotransmitter adrenaline from noradrenaline. In passing, we also noted that at the end of Chapter 7, the conversion of guanidinoacetate to creatine is a comparable amine *N*-methylation albeit in that case conversion of a secondary amine to a tertiary amine. 10 Analogously, histamine is *N*-methylated on N₃ of its imidazole ring as part of its metabolic breakdown in modulation of acute elevation of that bioactive amine during allergic responses.

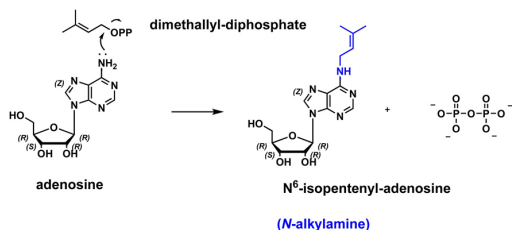
1. noradrenaline



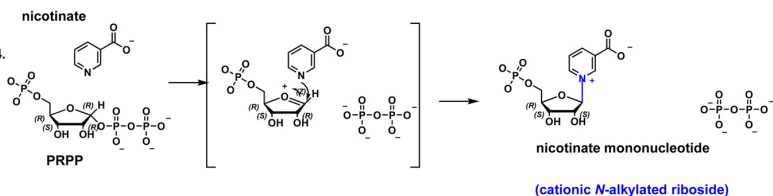
2.



3.



4.



Figures 8.5 Four of the six specific metabolic examples of amine *N*-alkylations:

noradrenaline *N*-methylation; phosphatidylethanolamine *N,N,N*-trimethylation; *N*-6-adenosine isoprenylation; nicotinate *N*-phosphoribosylation.

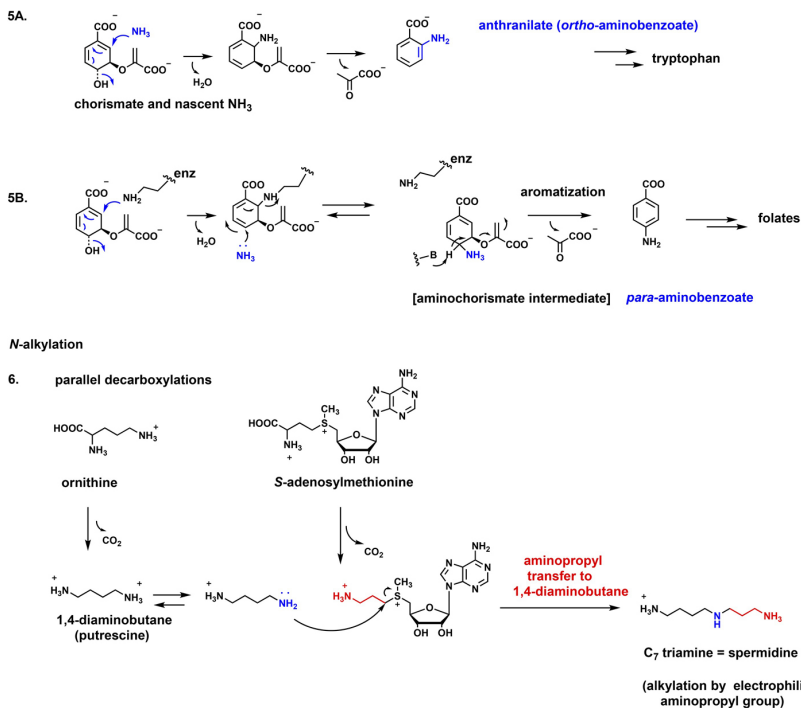


Figure 8.6 The remaining two of the six specific metabolic examples of amine *N*-alkylations: anthranilate and *para*-aminobenzoate aminations–decarboxylations; aminopropyl transfer from decarboxy-SAM.

The second line of [Figure 8.5](#) depicts three successive *N*-methyl transfers from SAM. The starting substrate is the membrane phospholipid phosphatidylethanolamine. (An alternative pathway is condensation of diacylglycerol with cytidine diphosphocholine.) During the three *ad seriatim* enzymatic methyl transfers, the amine in the ethanolamine moiety goes from primary to secondary to tertiary amine in the final product, phosphatidylcholine, which is a major lipid component of human cell membranes. [11](#) The choline moiety has a quaternized nitrogen with a permanent positive charge. It counteracts the single negative charge on the phosphodiester portion as a net neutral zwitterion. (We will take up *N*-trimethylation of histone protein side chain lysine nitrogens in the nitrogen proteomics Section IV, Chapter 16).

The third case in [Figure 8.5](#) is a five-carbon prenylation of the exocyclic N₆ amine of adenosine in plant metabolism. ¹² This reaction emphasizes the potential nucleophilicity of the adenosine exocyclic amine and is on the pathway to prenylated plant nucleoside phytohormones. This appears to be a straightforward capture of the isoprenyl C₁ carbocation ([Figure 8.4](#)) in the enzymatic transition state.

The fourth case is taken from the biosynthetic pathway to the nicotinamide adenine dinucleotide coenzyme forms of vitamin B₃. In Chapter 6 we noted previously that the final step in NAD⁺ biosynthesis was a glutaminase-mediated delivery of a nascent ammonia for the construction of the final step in side chain amide formation (Figure 6.3B). Here, we examine an earlier step in which the pyridine nitrogen (an aryl rather

than aliphatic nitrogen) of nicotinate (vitamin B₃) behaves as a nucleophile towards C₁ of cosubstrate 5-phospho-ribose-1-PP (PRPP). ¹³

In some analogy to the DMAPP early C₁-O cleavage to an allyl cation depicted in Figure 8.4, this features a comparable early C₁-OPP carbon-oxygen bond cleavage in PRPP reactions to generate an oxocarbenium ion across C₁-O₅. Capture of that electrophilic ribosyl C₁ by the pyridine nitrogen proceeds with inversion at C₁ to give the β -stereochemistry in the nicotinate mononucleotide product. Most consequential, the *N*-ribosylation imparts a *permanent positive charge* to the pyridine nitrogen, a pyridinium ion. (This positively charged pyridinium ion is the key to NAD⁺ acting as a hydride acceptor and NADH as a hydride donor as noted in Chapter 10).

Two examples of *N*-alkylations (among many) that show a broader underlying strategy are the microbial conversions of chorismite into *ortho*-aminobenzoate (anthranilate) and *para*-aminobenzoate (PABA) respectively (Figure 8.6). ¹⁴ Anthranilate is the immediate biosynthetic precursor to tryptophan while PABA is a key building block for folic acid, vitamin B₉.

The strategies of the two regiospecific syntheses are similar. Nascent ammonia is carved out of glutamine hydrolytically in each case by a glutaminase domain. For anthranilate, the nascent NH₃ is the competent nucleophile for the doubly vinylogous elimination that sets the incoming NH₂ *ortho* to the carboxylate that remains after subsequent pyruvate elimination and aromatization. To get to the *para*-aminobenzoate isomer, the initial nitrogen nucleophile is not nascent NH₃ but an enzymatic amine side chain, carrying out the same doubly vinylogous elimination of the 4-OH group. At that point the nascent NH₃ can attack C₄, reversing the doubly vinylogous addition step of the enzyme side chain. This tandem pair of eliminations places the NH₂ group from nascent ammonia at C₄. Now, aromatizing elimination of pyruvate yields PABA.

The anthranilate and PABA synthases use glutamine as carrier of latent NH₃ in its carboxamide group and release the reactive NH₃ *in situ*, perhaps to avoid indiscriminate reactivity with the sensitive dihydroaromatic ring system of chorismate in solution. The use of NH₃ or enzyme-NH₂-initiated single or iterated tandem doubly vinylogous elimination steps show a different kind of potential biological electrophile for *N*-alkylations. ¹⁴ Finally, the low energy, aromatizing elimination of the elements of pyruvate across C_{2,3} for anthranilate or C_{4,3} for PABA allow entry into arylamine frameworks. Anthranilate in particular is subsequently a source of polycyclic rings in microbial secondary metabolism.

The sixth and final example of Figure 8.6 involves the now familiar metabolites L-ornithine and S-adenosylmethionine. Each is a substrate for a particular amino acid decarboxylase, leading in parallel to a pair of amine products along with two molecules of CO₂. From ornithine, the 1,4-diaminobutane, *aka* putrescine, is produced, while SAM decarboxylation leads to the clumsily named descarboxy SAM. ¹⁵ One of the two primary amines of putrescine is directed to act as nucleophile towards C₃ of the aminopropyl chain of descarboxy SAM, claving the C₃-S⁺ linkage. The result is now a seven-carbon triamine named spermidine for its historical discovery in high concentrations in sperm. The coproduct is 5-methyladenosine. The triply positive spermidine molecules can act as charge neutralizing cations for electrostatic interaction with polyanionic DNA and create neutral stretches for genomic functional and structural effects.

8.1.4 Amine Glycosylations

A third common reaction type in primary metabolism is the transfer of glycosyl groups to cosubstrate nucleophiles. This is the central reaction type in oligosaccharides and polysaccharides (e.g. starch, cellulose) biosynthesis as well as in *N*-glycan chain assembly in glycoprotein formation. Figure 8.3 shows a prototypic reaction with UDP-

glucose as a paradigmatic donor of a glucosyl C₁-oxocarbenium ion equivalent to a nucleophilic substrate. For example, in glycogen biosynthesis that would be the C₄-OH of a terminal glucose residue in the growing glycogen chain (not shown).

An alternative form of activated glycosyl group, in the form of ribosyl C₁ electrophiles rather than hexosyl C₁ electrophiles, is provided by 5-phospho-ribosyl-1-pyrophosphate (PRPP). We have just noted the formation of nicotinate mononucleotide (NMN) during assembly of the coenzyme NAD⁺ in [Figure 8.5](#). Two additional phosphoribosylations from PRPP of nitrogen heterocycles are key chemical steps in the salvage pathway for conservation of nucleobases. ¹⁶ As shown in [Figure 8.7](#) in the active site of the relevant phosphoribosyl transferases, adenine reacts as a nucleophile at N₉ while uracil acts as a nucleophile at N₁, respectively, with PRPP to get back to the nucleotides AMP and UMP respectively. The regiospecificity of ribosylations accords with the basicity of the endocyclic N₉ and N₁, respectively, and salvages the heterocycles before they are eliminated and would need to be replaced by *de novo* heterocycle construction.

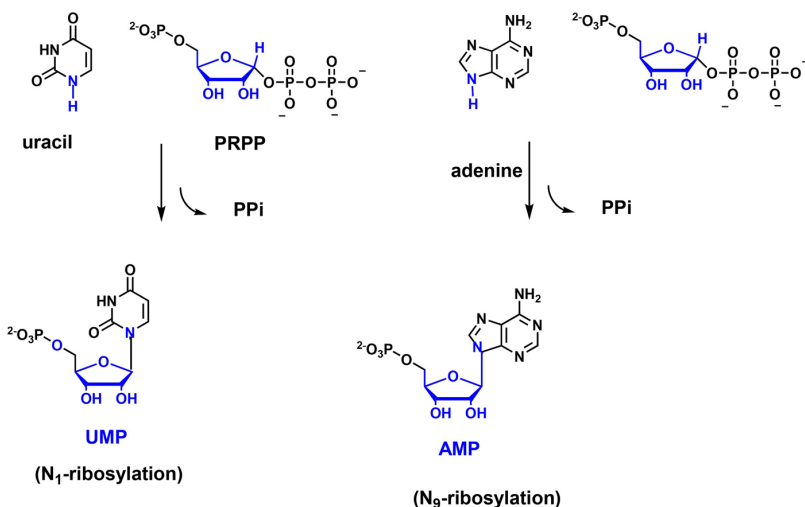


Figure 8.7 Salvage pathways for uracil and adenosine involve enzyme-catalyzed N-phosphoribosylations from 5-phospho-ribosyl-1-pyrophosphate (PRPP).

8.1.5 N-Sulfurylations and N-Phosphorylations

N-sulfurylation and N-phosphorylation reactions are plentiful in certain mammalian protein contexts, such as proteoglycan maturations and in the workings of the protein kinomes. The enzymatic transfer of the sulfonyl and phosphoryl electrophilic side chains of AMP-sulfate (APS) and phosphoadenosine phosphosulfate (PAPS), on the one hand, and from ATP, on the other hand, are substantially more restricted with low molecular weight amines.

One recent example of SO_3^- group transfer to an amine metabolite is found in enzymatic modulation of the shell fish-derived saxitoxin, a neurotoxin by virtue of ion channel blockade. ¹⁷ The structure of saxitoxin is itself intriguing, due to two guanidinium cationic groups (part of the active pharmacophore) being embedded within the heterocyclic core ([Figure 8.8A](#)). There is also a carbamate ester group, probably arising by carbamyl transfer from carbamyl phosphate. Two sulfonyl transfers occur from PAPS, one an N-sulfation, the other an O-sulfation. Notably, the first N-S

bond is formed as part of a sulfonamide moiety in gonyautoxin 5. This is an example of enzymes being able to activate a weak carbamate nitrogen nucleophile to attack the electrophilic sulfur in the mixed sulfuric–phosphoric anhydride side chain of PAPS.

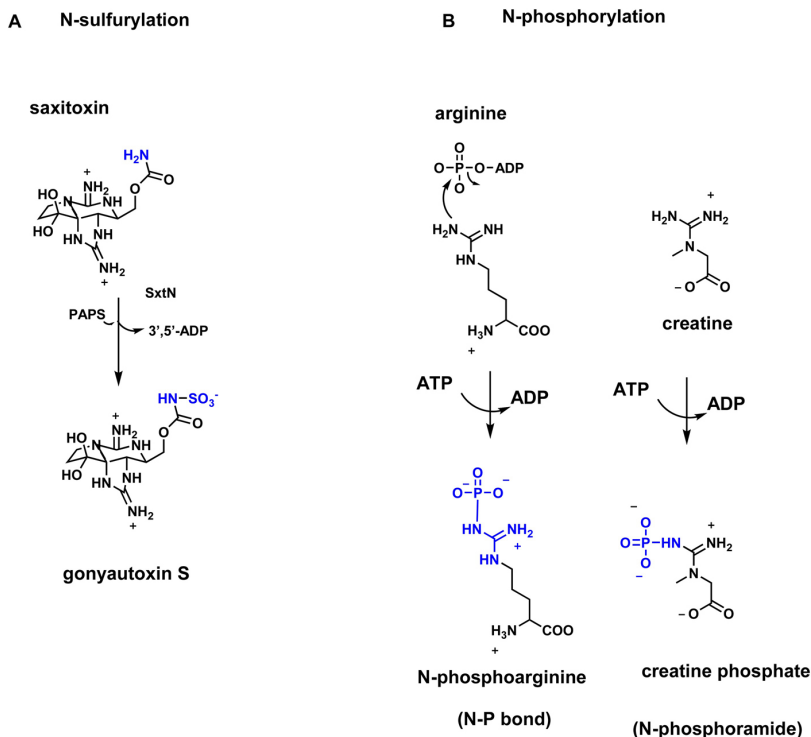


Figure 8.8 (A). *N*-sulfurylation of the saxitoxin scaffold; (B) *N*-phosphorylation of the guanidinium group in arginine or creatine.

N-phosphorylations follow a similar logic. The nitrogen nucleophile attacks the terminal P_γ of ATP, cleaving the P–O–P anhydride bond as a new N–P bond is produced in the phosphoramidate product. One might have anticipated that N_1 or N_3 of the imidazole nitrogen ring of histidine, lysine- N_6 , or ornithine N_5 would be recruited for such phosphoryl transfers. This is the case in nitrogen proteomics (Chapter 16), but to date these have not been seen as low molecular weight N– PO_3^{2-} metabolites. Perhaps they form, but themselves act as activated phosphoryl transfer reagents and so are not detected as accumulating metabolites.

Instead, one finds the *ab initio* unlikely *N*-phosphorylation of the guanidinium group of arginine and also of creatine (Figure 8.8B). The biggest surprise is that the pK_a values of the guanidinium groups are in the range of 13–14. Thus, at pH 7.5 there is one part in a million to one part in ten million of the deprotonated guanidine group in equilibrium with the great bulk of the positively charged guanidinium group with much lessened nitrogen nucleophilicity.

Nonetheless, the arginine kinases and creatine kinases, and related guanidinium kinases from other species, must be able to generate deprotonated nucleophilic guanidine moieties or their equivalents in active sites to serve as nitrogen nucleophiles. *N*-phosphoarginine accumulates in invertebrate muscle and *N*-phosphocreatine in vertebrate muscle, as phosphagens. 18 Phosphagens, as the name implies, serve as

phosphoryl transfer substrates to ADP to reform ATP in tissues with high demand from ATP (as in contracting muscle). They act as a kind of activated phosphoryl group battery.

8.1.6 *N*-Hydroxylations

We take up the subject of *N*-hydroxylations and one-electron vs. two-electron nitrogen chemistry in Chapters 12 and 13. We simply note here that enzymes exist in bacteria that make N₆-OH-lysine and that enzymes in humans and other vertebrates make N-OH-arginine (Figure 8.9). We will examine the mechanisms of both enzymes in those later chapters. The *N*-hydroxylysine residues get incorporated into siderophore frameworks for Fe³⁺ ion chelation and import. The N-OH-arginine is an intermediate in the various isoenzyme forms of human nitric oxide synthases.

Two different cofactor types for nitrogen monooxygenases deliver [OH⁺] or [OH[•]] equivalents

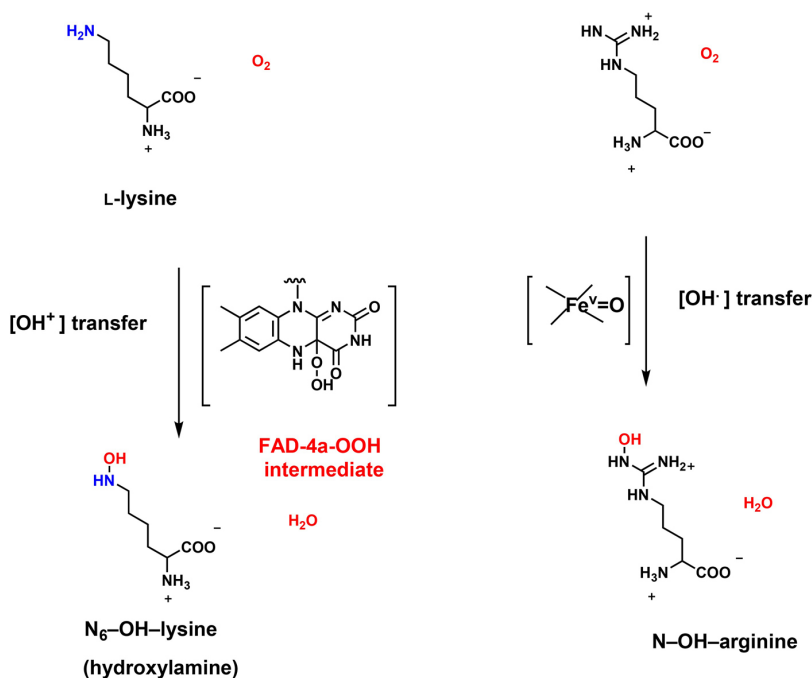


Figure 8.9 Flavoenzyme oxygenases and heme-iron *N*-oxygenases: [OH⁺] vs. [OH[•]] transfers.

Two types of oxygen reactive cofactors are employed in the two enzymes and reflect different oxygen transfer mechanism. The lysine N₆ monooxygenase utilizes FAD, a riboflavin-based coenzyme to transfer one oxygen to the N₆-amine with the polarity of an [OH⁺] equivalent as the flavin-OOH adduct decomposes. In contrast, the oxygen transfer reagent in iterated oxygenations by the nitric oxide synthases is a high valency porphyrin iron oxide species that employs radical chemistry in transfer of an [OH[•]] equivalent to a nitrogen radical. The first but not the second oxygenase fits into the

rubric of transfer of an electrophilic substrate fragment to a cosubstrate nitrogen nucleophile.

8.2 Imine–Iminium Ion Reactions to Form or Cleave C–C Bonds

In Chapter 6 we noted that iminium formation and reduction to the amino acid glutamate was a prime route for fixation of NH_3 into keto acid scaffolds as it moved from inorganic to organic metabolism. [Figure 8.10A](#), a partial reprise of [Figure 8.1](#), explicitly reminds us that the nitrogen in amines (and in parent NH_3) is nucleophilic.

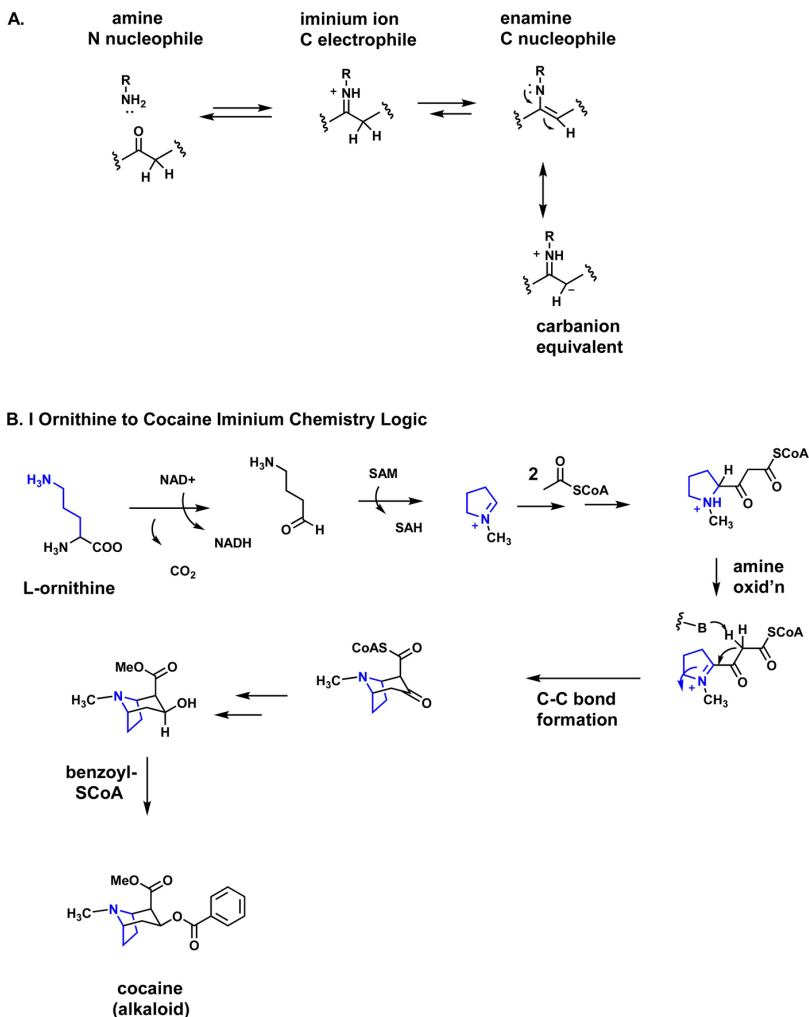


Figure 8.10 Iminium ion–enamine tautomers in C–C bond formations: (A) nitrogen nucleophiles vs. carbon electrophiles and carbon nucleophiles; (B) ornithine to cocaine: iminium ion chemical logic in biosynthesis.

The ready and reversible condensative dehydration with aldehydes or ketones to form iminium ions, dampens down the nitrogen nucleophilicity. On the other hand, the nitrogen in an imine is much more basic than the oxygen in the starting carbonyl, so the imine is readily protonated under enzymatic catalytic conditions. The carbonyl carbons in iminium ions are thus strongly electrophilic centers. The mechanism of plant-based conversion of L-ornithine to cocaine in [Figure 8.10B](#) illustrates the role of iminium ions in C–C bond forming steps. ¹⁹

Finally, biology takes advantage of iminium ions in which one of the adjacent carbon atoms has a hydrogen substituent, as illustrated in [Figure 8.10A](#). Abstraction of one of those hydrogen atoms as a proton is a low energy C–H cleavage step because the resultant carbanion negative charge is delocalized into the nitrogen as the *enamine* resonance contributor. Because nitrogen is more basic than carbon, the predominant contributor is the enamine structure. Thus, one now has an accessible low energy carbanion equivalent as a result of iminium ion formation and facile tautomerization to the enamine.

We take a brief look at two examples in plant alkaloid biosynthesis to illustrate first the iminium ions as electrophiles and then enamines as carbanion equivalents. There are many more examples to be found in alkaloid biosynthesis, but they are beyond the scope of this volume about nitrogen chemical biology.

8.2.1 Ornithine to Cocaine: Four-iminium Intermediates

Coca plants convert the five-carbon amino acid ornithine, a nonproteinogenic but still metabolically central amino acid, into the bicyclic ring system of tropane alkaloids including the psychoactive cocaine ([Figure 8.10B](#)). The pathway utilizes ornithine and two molecules of acetyl CoA to give one molecule of CO₂ and the eight-carbon tropane scaffold. As was the case in the spermidine biosynthetic pathway in [Figure 8.6](#), the first step is catalyzed by PLP-containing ornithine decarboxylase, accounting for the one CO₂ molecule as it produces 1,4-diaminobutane. An NAD⁺-dependent oxidoreductase takes one of the diaminobutane amines to the imine. Nonenzymatic hydrolysis of the imine releases NH₃ and the aminobutanal. That can undergo intramolecular cyclization to the δ -1-pyrroline, the cyclic five-membered imine in prototropic equilibrium with the iminium ion shown. The neutral imine form (nitrogen as nucleophile) is then methylated on nitrogen by a [CH₃⁺] equivalent from S-adenosylmethionine.

The resulting cationic N-methyl iminium ion is a good electron sink for iterated condensation with acetyl CoA molecules. ¹⁹ The first C–C bond forms as the accessible C₂ carbanion of acetyl-CoA (think thioester enolate) attacks the pyrroline C₁-iminium ion. A second carbanion addition of an acetyl CoA-derived acetyl unit generates the 3-ketobutyryl CoA with an attached N-methylpyrrolidine ring.

Now, another NAD⁺-dependent oxidoreductase recreates the cyclic imine as electron sink, again by dehydrogenation. The C₂ methylene carbon has acidic hydrogens because the resulting carbanion is delocalized and stabilized both as the enolate at C₂ and as the thioester enolate at C₁. That C₂ carbanion adds *intramolecularly* into the C₁ of the pyrrolidinium ring, creating a new C–C bond that forms the fused 5,5- bicyclic tropane framework. To go from the N-methyltropane–CoA thioester to cocaine involves thioester hydrolysis, O-methyl ester formation, stereospecific reduction of the tropane ring carbonyl to the β -alcohol and esterification with the benzoyl moiety from benzoyl CoA. ⁴

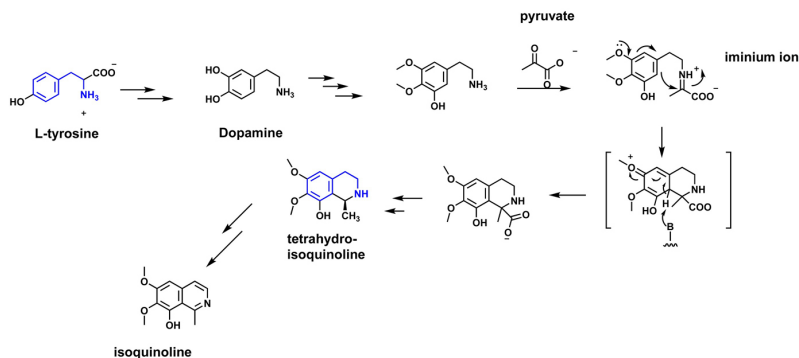
Four iminium ions are formed along the way, the first as nascent product from ornithine decarboxylase action, the second in the oxidoreductase reaction of 1,4-diaminobutane to aminobutanal. The third one forms the cyclic pyrrolidinium ion to set up C–C bond formation to attach the acetoacetyl side chain. The fourth iminium ion is the electron sink for tropane scaffold formation. The cocaine pathway shows the

versatility of iminium ions as electron sinks for C–C bond formations.

8.2.2 Tyrosine to Tetrahydroisoquinoline Alkaloid Framework

Plants turn several common amino acid scaffolds into a wide range of heterocyclic alkaloid frameworks. One simple pathway starts with tyrosine and proceeds through dopamine. The three-carbon keto acid metabolite pyruvate is then a cosubstrate for iminium ion chemistry that sets the bicyclic tetrahydroisoquinoline [20](#) scaffold as summarized in [Figure 8.11](#).

A. Iminium Chemistry In the Conversion of Tyrosine and Pyruvate to Tetrahydroisoquinoline Alkaloids



B. Indole Ring of Tryptophan engages in Imine/Enamine Tautomerization

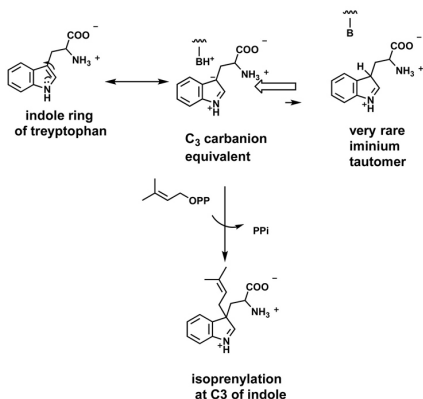


Figure 8.11 Iminium ion–enamine tautomers in C–C bond formations: (A) iminium ions in condensation of tyrosine and pyruvate on the way to tetrahydroisoquinoline alkaloids; (B) prenylation of the indole ring of tryptophan at carbon 3.

Tyrosine is converted to dopamine both in animals (where it is a central nervous system transmitter) and in plants (where it is an alkaloid precursor). Two enzymatic steps, *ortho* hydroxylation on the phenol ring and then PLP-mediated decarboxylase action, generate dopamine (precursor also to the noradrenaline illustrated in [Figure 8.5](#)). Dopamine is then *bis*-methylated at the 3,4-diol moiety and oxygenated at C₅ of the ring. The next enzyme adds pyruvate into the primary amine of the *bis*-

methoxyphenolic amine to create the indicated iminium ion as an electron sink. Using the iminium terminus, a new C–C bond forms with relay participation from the 3-methoxy group. Rearomatization yields the bicyclic scaffold of the tetrahydroisoquinoline framework, a starting point for multiple downstream enzyme manipulations, including double dehydrogenations to the fully aromatic isoquinoline alkaloid family.

8.3 Enamine Intermediates as Carbanion Equivalents for C–C Bond Formation

The proteinogenic amino acid tryptophan uniquely contains a bicyclic indole side chain. We will examine indole reactivity in Chapter 9, in the context of its heterocyclic chemistry in plant and fungal conditional metabolism. Here, we introduce the fact that free tryptophan expresses its chemical ability for enamine chemistry by utilizing C₃ of the indolic pyrrole ring as a carbon nucleophile [21](#) ([Figure 8.11B](#)).

Parsing the N₁, C₂, C₃ part structure of the indole ring indicates that this part of the pyrrole is an enamine structure. The lone pair at N₁ can participate in a minor resonance contributor that has carbanion density at C₃ and an iminium group at C₂–N₁ as depicted in [Figure 8.11B](#). This accounts for the long-known reactivity of indoles at C₃ as carbon nucleophiles.

A common example of that reactivity is observed in plant secondary metabolism when prenyl transferase enzymes add an electrophilic prenyl group to C₃ of the indole ring of tryptophan or its decarboxylation product tryptamine. [21](#) Early dissociation of the dimethylallyl pyrophosphate–C₁–O bond gives an allyl cation transition state. Capture of that stabilized carbocation at C₁ by the C₃ carbanion equivalent of tryptophan yields the indicated C₃-prenylated adduct. Note that the indole ring has been trapped as the C₂–N₁ iminium form. We will see in the next chapter, on nitrogen heterocycles, that iminium ion in turn sets up trapping reactions that form new bonds, and thereby new ring systems, fused to the pyrrole ring of the indole framework.

References

1. F. M. Raushel, J. Thoden and H. M. Holden The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia, *Biochemistry*, 1999, **38**, 7891-7899.
2. R. J. Ouellette and D. Rawn, *Organic Chemistry, Key Concepts, Problems and Solutions*, Elsevier, 2015.
3. C. T. Walsh, B. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
4. C. T. Walsh and Y. Tang, *Biosynthesis of Natural Products*, Royal Society of Chemistry Press, London UK, 2017.
5. L. Caldovic, N. Ah Mew, D. Shi, H. Morizono, M. Yudkoff and M. Tuchman N-Acetylglutamate synthase: structure, function and defects, *Mol. Genet. Metab.*, 2010, **100**, S13-S19.
6. A. Kumari, P. Pasini, S. K. Deo, D. Flomenhoft, S. Shashidar and S. Dauert Biosensing systems for the detection of bacterial quorum sensing molecules, *Anal. Chem.*, 2006, **78**, 7603-7609.
7. J. E. Gouaux and W. N. Lipscomb Three dimensional structure of carbamoyl phosphate and succinate bound to aspartate transcarbamoylase, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 4205-4208.

8. J. B. Walker Creatine biosynthesis, regulation and function, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1979, **50**, 177-242.
9. M. E. Fontcave, M. Atta and E. Mulliez S-adenosylmethionine: nothing goes to waste, *Trends Biochem. Sci.*, 2004, **29**, 243-249.
10. S. Stockler, I. D. Brandt and F. Hanefeld Guanidinoacetate methyltransferase deficiency: the first inborn error of creatine metabolism in man, *Am. J. Hum. Genet.*, 1996, **58**, 914-922.
11. Y. Philip, *The Membranes of Cells*, London, 2016.
12. M. Laloue, C. Terrine and J. Guern Cytokinins: metabolism and biologic activity of N⁶-isopentenyladenosine and N⁶-isopentenyladenine in tobacco cells and callus, *Plant Physiol.*, 1977, **59**, 478-483.
13. N. G. Dulyaninov, E. Podlepa, L. V. Toilokhanova and V. Y. Bykhovsky Salvage pathway for NAD biosynthesis in *Brevibacterium ammoniagenes*: regulatory properties of triphosphate-dependent nicotinate phosphoribosyltransferase, *Biochim. Biophys. Acta*, 2000, **1478**, 211-220.
14. C. T. Walsh, S. Haynes and B. Ames Aminobenzoates as building blocks for natural product assembly lines, *Nat. Prod. Rep.*, 2011, **29**, 37-59.
15. Y. Ikeguchi, M. Bewley and A. E. Pegg Aminopropyltransferases: function, structure, genetics, *J. Biochem.*, 2006, **139**, 1-9.
16. S. C. Sinha and J. Smith The PRT protein family, *Curr. Opin. Struct. Biol.*, 2001, **11**, 733-739.
17. A. L. Lukowski, N. Denomme, M. E. Hinze, S. Hall, L. L. Isom and A. R. H. Nrayan Biocatalytic detoxification of paralytic shellfish toxins, *ACS Chem. Biol.*, 2019, **14**, 941-948.
18. W. R. Ellington Evolution and physiologic roles of phosphagens, *Annu. Rev. Physiol.*, 2001, **63**, 289-325.
19. R. J. Robins, T. Abraham, J. Parr, J. Eagles and N. J. Walton The biosynthesis of alkaloids in *Datura stramonium*: the identity of the intermediates between N-methylpyrrolidium salt and tropinone, *J. Am. Chem. Soc.*, 1997, **119**, 10929-10934.
20. A. Mirchison Saturated nitrogen heterocycles, *J. Chem. Soc., Perkin Trans. 1*, 2000, **1**, 2862-2892.
21. W. Xu, D. Gavia and Y. Tang Biosynthesis of fungal indole alkaloids, *Nat. Prod. Rep.*, 2014, **31**, 1474-1487.

Nitrogen Heterocycles

9.1 Introduction to Nitrogen Heterocycles

One could argue that organic nitrogen heterocyclic chemistry is the apotheosis of nitrogen chemical biology. The purine and pyrimidine nucleobases are the building blocks of RNA and DNA for information storage and transfer. The indole ring of tryptophan is by far the most versatile side chain of any of the 20 proteinogenic amino acids. Nitrogen-based chemistry of pyrroles enables tetrapyrrole macrocycle assembly to form the pigments of life. Seven of the eight B vitamins function as heterocyclic coenzymes in almost every major metabolic pathway. A remarkable array of nitrogen heterocycles, these heteroaromatic ring systems enable the chemistry of life.

One could categorize nitrogen heterocycles by ring size, by the number of nitrogens and/or other heteroatoms, by physiological functional niches. ¹ Any one of the topics in the preceding paragraph can easily fill more than one volume of this size. In this chapter we take up selected illustrative examples, with acknowledged idiosyncrasy in favor of illustrating a range of biologys.

Figure 9.1A Shows the structures of the three proteinogenic amino acids that contain heterocyclic rings as well as the two pairs of pyrimidine and purine nitrogen heterocycle building blocks for RNA. The pyrimidine 5-methyluracil (thymine; not shown) replaces uracil in DNA. Each of these seven nitrogen heterocycles offers different chemical biology.

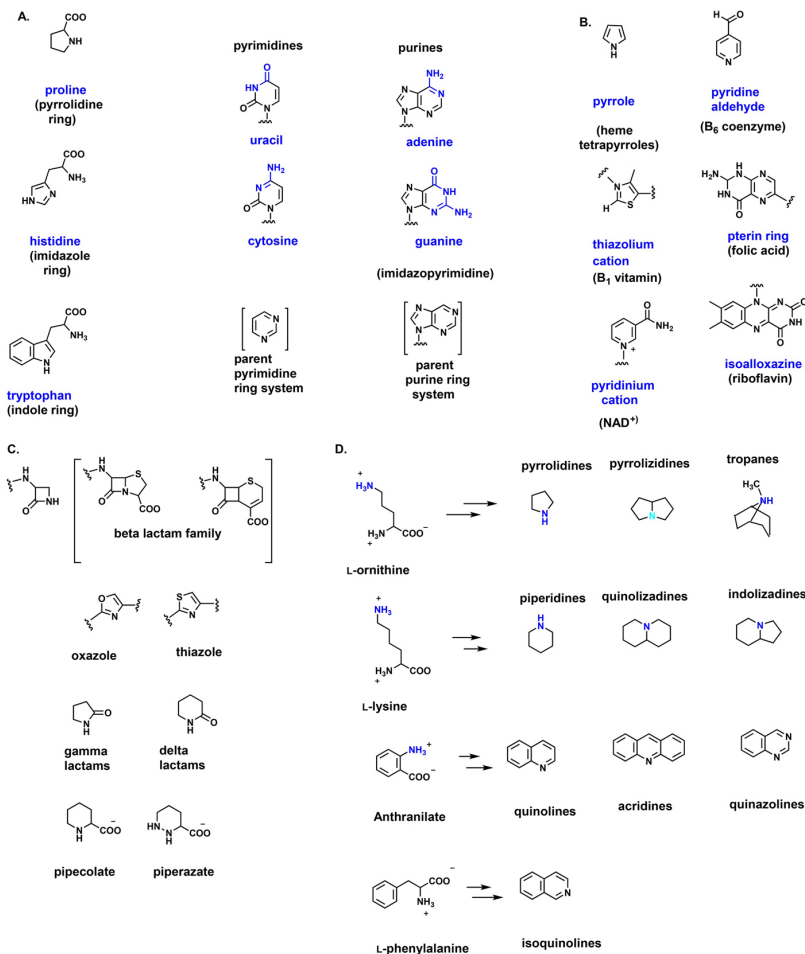


Figure 9.1 Nitrogen heterocycles in metabolism. (A) Three proteinogenic amino acid heterocycles and purine and pyrimidine building blocks for nucleic acids. (B) Additional nitrogen heterocycles in primary metabolism. (C) Selected nitrogen heterocycles in secondary metabolism. (D) Classes of heterocyclic alkaloids derived from ornithine, lysine, phenylalanine and anthranilate (beta-amino acid).

All three of the amino acid side chain heterocycles in [Figure 9.1](#) contain a five-membered heterocycle with one or two nitrogen atoms. In proline the parent ring is a pyrrolidine. It confers no special reactivity properties to proline other than the conformational constraints from locking the nitrogen as a secondary amine in the ring structure.

The imidazole ring of histidine is a much more formidable functional group. The pK_a between the doubly protonated imidazolium ring and the monoprotonated imidazole ring, ([Figure 1.2B](#)) is 6.0. Thus, at pH 7 about 90% of histidine will have the neutral imidazole ring in which the lone electron pair of the sp^3 N-H and sp^2 (C=N) group are available as both a base and a nucleophile. This means a higher mole fraction of free base or nucleophilic nitrogen in the imidazole side chain than any of the alpha amino groups by a factor of 10^3 to 10^4 , since their pK_a values cluster between 9 and 10.

The indole side chain of tryptophan has an embedded pyrrole ring fused to a phenyl

ring. We noted in Chapter 8 that the resting state of the pyrrole moiety contained an enamine moiety that can act as C₃ nucleophile. The pK_a for protonation of the indole nitrogen occurs only in strongly acidic solution, far beyond the physiological range. Only the neutral indole N–H figures in its chemical biology.

Figure 9.1B illustrates six additional nitrogen heterocyclic motifs that play central roles in primary metabolism. We noted the pyrrole moiety fused to a benzene ring in the indole ring of tryptophan, but the monocyclic pyrrole is at the center of the tetramerization pathways to the heme, chlorophyll, vitamin B₁₂, and methanogen F₄₃₀ tetrapyrrole macrocycles. We examine pyrrole chemical biology in detail in Chapter 11. The other five nitrogen heterocycles in Figure 9.1B, thiazolium cation, pyridinium cation, pyridine aldehyde, bicyclic pterin, and tricyclic isoalloxazine, are all the business ends of coenzyme forms of five of the B vitamins that are examined in Chapter 10.

Figure 9.1C illustrates some additional heterocycles including β, γ, and δ-lactams (cyclic amides). We will look briefly at the β-lactam class of antibiotics, including the penicillins and cephalosporins, as the bestselling class of antibiotics globally over the past 80 years. Finally, plants and fungi make hundreds to thousands of secondary natural products by conditional metabolic pathways, including a plethora of heterocycles with one or more basic nitrogen atoms, collected historically under the generic category of alkaloids (Figure 9.1D). Natural product biosynthesis in general and alkaloid biosynthesis in particular is not the topic of this volume (readers are directed to Walsh and Tang 2 for an introduction to the molecular logic of complex ring assemblies). We note only in passing here that four of the amino acids we have encountered thus far [three α-amino acids and the β-amino acid anthranilate (see Chapter 8 for installation of the anthranilate amino group into the framework of chorismate)] are starting points for specific nitrogen heterocyclic alkaloid classes.

9.2 Amides, Amidines and Guanidines as H-bond Participants in DNA Base Pairs

Alongside proteins, the role of nitrogens in the five nucleotide bases that build RNAs and DNAs in all cells may be the most obviously consequential in biological information transfer. The Watson–Crick scheme 3 of base pairing that holds the two strands of DNA together in double helical array depends on the inter-base hydrogen bonds between A–T and C–G base pairs. A–T pairs have two hydrogen bonds while C–G pairs have three hydrogen bonds (Figure 9.2A), reflected in the lower melting temperature of A–T rich runs in DNA.

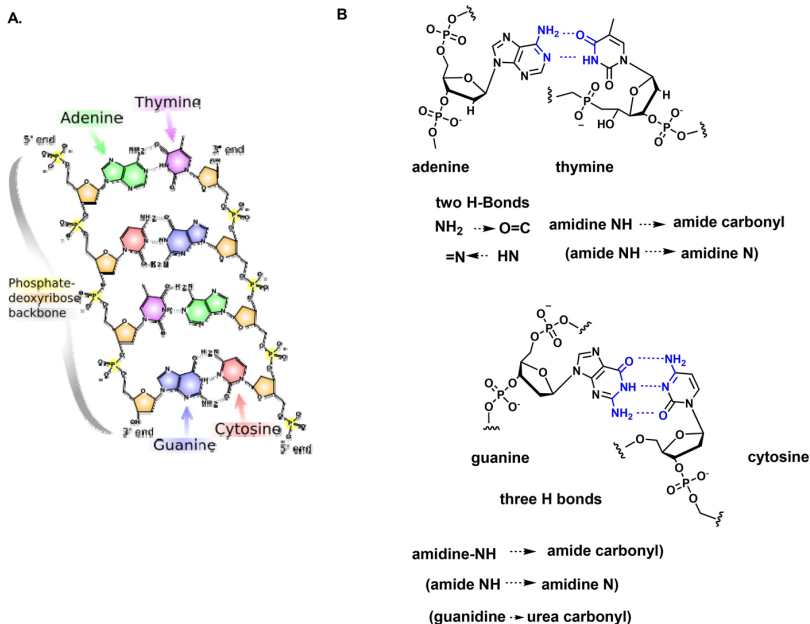


Figure 9.2 Hydrogen bonding between nitrogen heterocyclic building blocks in DNA. (A) View of a tetranucleotide region in double stranded DNA. Reproduced from https://en.wikipedia.org/wiki/DNA#/media/File:DNA_chemical_structure.svg, under the terms of the CC BY-SA 3.0 license <http://creativecommons.org/licenses/by-sa/3.0/>, credit: Madeleine Price Ball. (B) A-T and G-C pairs with amide, amidine, urea and guanidine H bonding groups identified.

In a canonical A-T base pair there is a directional N-H $\cdots$ O base pair (amine to amide carbonyl) and an amine to amidine (N $\cdots$ NH) base pair (Figure 9.2B). The N-H $\cdots$ O hydrogen bond is from N₆ of the adenine to the C₄ carbonyl oxygen of thymine on the opposite DNA strand. The polarity of the second H-bond in the A-T pair is opposite, with N₃-H of thymine as the donor to N₃ of adenine as acceptor.

The C-G base pairs have two H bonds where the guanine is the donor and one where it is the acceptor (Figure 9.2B). The exocyclic N₂-H and the N₃-H of guanine (parts of the same guanidine moiety in the GMP residue) are the two H-bond donors. The acceptor atoms on the paired cytosine are, respectively, the ureido carbonyl oxygen O at C₂ and the amidine nitrogen N₃ in cytosine on the opposite strand. The third H bond has the opposite polarity with the amidine C₄-NH₂ of cytosine as donor to the amide carbonyl O₄ of the guanine. (Thus, *amidines* can function as either H-bond donor or H-bond acceptor, depending on tautomer distribution and partner base pair.)

A key set of observations is that the hydrogen bonds so crucial in maintaining the DNA double helix in every genome involve *amidine*-, *guanidine*- and *amide nitrogens* as H bond donors. The H-bond acceptors in the bidirectional A-T and G-C pairs are *amide carbonyls*, *urea carbonyls* and *amidine nitrogens*. These are clearly not electron rich simple amine groups as H-bond donors. They emphasize again the central role that amidines, ureas and guanidine groups play both in metabolism (e.g. Chapter 7) and as the structural elements in the nitrogen heterocyclic pyrimidine and purine base pairs. Much of the stability of DNA double helices rests on these N-H $\cdots$ N amide-amidine hydrogen bonds, and thereby on the ordering of basicity among nitrogens in amine, amide and amidine functional groups.

This simplified, first order, analysis of amine, amide and amidine functional groups in purine and pyrimidine base pairs raises the question: how do they form? What are the origins of the nitrogen atoms as purine and pyrimidine building blocks of RNA and DNA are biosynthesized?

9.3 Parsing the Origin of Nitrogens in Pyrimidines and Purines Reveals Simple Assembly Logic

The four or five nucleobases that are the information coding units for RNA and DNA fall into the pyrimidine and purine ring categories (Figure 9.1). The placement of oxygen and nitrogen functionalities within the pyrimidine and purine ring system are crucial functional groups for determining specificity, e.g. of A–U (in RNA), A–T and G–C pairs and thereby information for replication, transcription and translation of the coding information in genomic DNA and in messenger RNAs.

The location of the H-bonding donor and acceptor nitrogen and oxygen atoms in pyrimidines and purines arise from the simple biosynthetic logic that organisms devised (early on in evolution) for the biosynthesis of these key nitrogen heterocyclic building blocks. Examination of the origin of the two (uracil) to three (cytidine) nitrogens in the pyrimidines and the five nitrogens in adenine and guanine nucleotides illuminates the logic of ring assemblies that use very small and common metabolites as building blocks by very simple reaction mechanisms, consistent with early (rudimentary?) assembly of these stable heterocycles.

The reaction mechanisms are consistent with the types of *N*-acylation reactions we have summarized in Chapters 6–8. We start with the single ring pyrimidine and then turn for comparative logic to the bicyclic purines, noting that the bicyclic purine is a fused imidazolo-pyrimidine system.

9.3.1 Two Nitrogens in Uridine, Three in Cytidine

The six-membered cyclic framework of pyrimidines is created in only two enzymatic steps from the common pool of metabolites L-aspartate and carbamyl phosphate 4 (Figure 9.3). Aspartate transcarbamylase deprotonates the major ammonium ion form of aspartate in its active site to generate the nitrogen nucleophile to attack the electrophilic acyl phosphate, carbamoyl phosphate, on its carbonyl group. As the enzyme name states, this effects a net carbamyl transfer to the attacking amine. The product *N*-carbamyl aspartate has a newly formed *ureido* group, as is the case for any carbamyl transfer to an amine. The carbamyl aspartate is the first committed metabolite, sending flux to pyrimidine formation billions of times per cell cycle to create enough new DNA building blocks for full genome replication and for all the ribosomal, transfer and messenger RNAs (rRNAs, tRNAs and mRNAs) transcribed from the genome.

NH₃ via glutamine. (C) Facile tautomeric equilibria in amidines (cytosine and adenine rings) and in guanidines (guanine ring).

The second enzyme, dihydroorotase, catalyzes the dehydrative cyclization of the acyclic *N*-carbamyl aspartate to the cyclic dihydroorotate. A new amide bond has formed between the carbamate amido-nitrogen as a nucleophile and the C₄ carboxylate of the aspartate moiety as an electrophile. There are two unusual features to this amide formation cyclization. First, the carbamyl nitrogen as part of a ureido group is at best a weak nucleophile and likely to be sluggish as a reactant. Second, this is a *rare case* of enzymatic amide formation where the carboxylate component has not been converted to an electrophilic acyl phosphate, acyl-AMP, or acyl thioester. The OH[−] group, rather than OPO₃^{2−}, AMPO[−] and CoAS[−], is a higher energy leaving group.

Nonetheless, the zinc-dependent dihydroorotase manages cyclizing amide formation readily, perhaps because of high intramolecularity of the two reacting groups. This amide bond has created the cyclic dihydro-pyrimidine framework. Note the source of the two nitrogens: ammonia (*via* carbamyl phosphate) and the aspartate amino group. The full scaffold of aspartate, including the three contiguous carbons 2, 3 and 4 of aspartate have been incorporated. They form the three-carbon backbone in the bis-nitrogen heterocycle with its ureido and amide carbonyl part structures.

The third enzyme is a dehydrogenase converting the dihydroorotate to the heteroaromatic ring system of orotate. The shape of the pyrimidine bases is now in full view. The fourth enzyme fits in the category of an *N*-ribosyltransferase 5 of the type profiled in Figure 6.3. The C₁–O₅-oxocarbenium ion in a transition state arising from early C₁–OPP cleavage of bound phosphoribosyl pyrophosphate (PRPP) substrate gives a C₁ ribosyl electrophile to be captured by N₁ of orotate. This N–C bond formation also relies on an attenuated nitrogen nucleophile since N₁ of orotate is part of a ureido group. The resultant orotidine monophosphate (OMP) is an enzyme-catalyzed decarboxylation away from UMP.

The third nitrogen, present in cytidine nucleotides but not uridine nucleotides, occurs at the very last step in conversion of UTP to CTP, an amide to amidine (imine derivative of an amide) enzymatic reaction catalyzed by CTP synthetase 6 (Figure 9.3B). As noted in Chapter 6, the key issue is to reverse the polarity of the C₄ amide carbonyl of UTP to become a low energy leaving group. Trapping of the minor isoamide anion form by formation of the isoamide–AMP ester, sets up attack of nascent NH₃ derived from glutaminase domain action.

The tautomeric forms available in both the amidines of cytosine and adenine bases are shown in Figure 9.3C. Also displayed is the comparable guanidine tautomer for guanidine, emphasizing the utility of the tautomeric forms in altering hydrogen bonding patterns.

Thus, summing up the origin of the three nitrogens in cytidine nucleotides, the first and the third come from NH₃ by way of two different intermediates, carbamoyl phosphate vs. glutamine's carboxamide. The second nitrogen is from aspartate. Carbons 2, 3 and 4 of the aspartate framework are brought along and become carbons 6, 5 and 4 in the uridine and cytosine end products. The use of aspartate as the amino group donor has parallels with, and contrasts to, the use of aspartate as the amino donor in purine assembly, considered next.

9.3.2 Five Nitrogens in AMP and GMP

By comparison to the monocyclic pyrimidine nucleotides, biogenesis of the purines AMP and GMP is longer in steps but equally simple, albeit variant in strategy 7 (Figure 9.4). First, in contrast to pyrimidine nucleotides where phosphoribosylation occurs at step four in the five-step biosynthetic pathway to UMP, the first step in purine

formation is phosphoribosylation of nascent ammonia [released by glutaminase action (Chapter 6) to give 5-phospho-ribosyl-1-amine (*N*-ribosylation reaction)]. The purine bis-heterocycle is then built from the ground up.

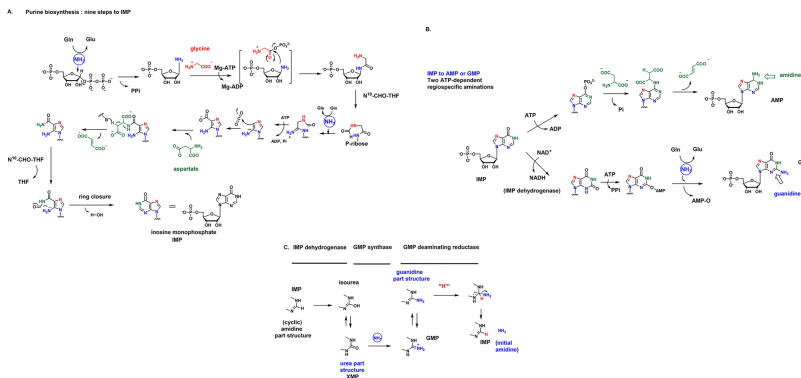


Figure 9.4 Three aspects of purine metabolism. (A) A nine-step assembly process to IMP. (B) Distinctive amination strategies from IMP to AMP vs. IMP to GMP. (C) Three enzymes in the cycle from IMP to GMP and back to IMP (amidine to urea to guanidine to amidine).

In analogy to the use of the full skeleton of the proteinogenic amino acid aspartate in pyrimidine assembly, the second enzyme of the purine pathway incorporates glycine intact to give glycinamide ribotide (GAR). The 5-phosphoribosyl-1-amine-NH₂ is the nucleophile, forming an amide bond to the glycine carboxylate. As anticipated, the glycine carboxylate has to be activated as an electrophile to undergo attack by the phosphoribosylamine (PRA) amine group. That is the role for ATP, formation of a glycy phosphate mixed anhydride that gets captured by PRA to form the amide linkage in GAR. Thus, the second nitrogen in comes from glycine.

Next is delivery of a one-carbon unit at the formyl oxidation state, presented by N¹⁰-formyl tetrahydrofolate, a carrier coenzyme discussed in Chapter 10. The NH₂ group of the glycy moiety of GAR attacks that formyl group to generate the second amide linkage in *N*-formyl-GAR (FGAR). The fourth enzyme contains another glutaminase domain, ⁶ releasing nascent ammonia that is tied up by the glycinamide carbonyl as the imine. To this point three nitrogens have been added as the imidazole ring is close to forming, two from nascent NH₃ by a pair of glutaminase domains and one from glycine in acyclic FGAR.

Dehydrative cyclization by the fifth enzyme creates the aminoimidazole ribotide framework (AIR). The sixth enzyme adds a second one-carbon unit, this time from CO₂ in an unusual carboxylation to get to the amino and carboxy substituted imidazole ribotide. At this juncture the fourth nitrogen is added by way of the amino group of aspartate, forming an amide to the carboxyl group substituent on the imidazole. This is the third amide bond formed.

The product succinyl-aminoimidazole carboxamide ribotide is reminiscent of the incorporation of the full aspartate scaffold at the start of pyrimidine biosynthesis by action of aspartate transcarbamylase (Figure 9.3). However, the next enzyme is one of those lyases that cleaves out the four-carbon scaffold of the aspartyl moiety as fumarate while leaving just the amino group of aspartate as the carboxamide. The growing chain is one carbon short of a purine. That one carbon is provided in a reprise of the earlier *N*-formylation via N¹⁰-formyl-tetrahydrofolate (THF). That is the fourth amide bond forming step in the pathway. One last, ring closing step is attack of what

becomes N₁ of the pyrimidine ring part of IMP on the newly added formyl group (C₂ in the final purine). Loss of water gives the pathway end product, inosine monophosphate.

Of the four amide bonds formed during stepwise imidazole and then fused pyrimidine ring formation, one persists in IMP. The others have undergone dehydration to the three endocyclic C=N double bonds in IMP. The ribosyl group is appended to the nitrogen at the bottom of the imidazole ring (N₉) and has been on that nitrogen from the very first step. Of the four nitrogens added as nucleophiles, two have come from nascent NH₃ by way of *in situ* glutaminase actions, one from glycine, and one from aspartate in the two-step aspartate addition–fumarate elimination reaction sequence. Of the five carbons in the purine bicycle, two come from glycine, two from formate *via* formyl-THF (ultimately that formyl group derives from C₃ of serine) and one from CO₂. This logic is almost the equivalent of “earth, air, fire and water” in its chemical simplicity, consistent with an early, even prebiotic assembly route to this key heterocycle.

IMP is not itself one of the two purine metabolites in RNA and DNA. Rather, it is the branch point for a one step amination to AMP, or a two-step [*via* xanthosine monophosphate (XMP)] dehydrogenase, aminase reaction sequence to GMP (Figure 9.4B). The IMP dehydrogenase transfers a C₂-hydride from IMP and adds in a water molecule to give an initial hydroxy tautomer of XMP that ketonizes to the predominant tautomer. That XMP tautomer has a ureido part structure at N₁–C₂–N₃. Its subsequent amination, as always, converts a *urea* to a *guanidine* in GMP. Separately, the direct conversion of the amide carbonyl of IMP to the exocyclic amine of AMP by the AMP synthase pathway converts the *amide grouping* to the *amidine*, as emphasized in Chapter 8.

The enzymatic conversion of IMP to GMP and AMP shows the interrelationships between amides and amidines as well as between ureas and guanidine groups. The amidine and guanidine groups have lower energy barriers than the amide and urea groups for populating the less favored tautomers (Figure 9.3C), a consequence of the replacement of each oxygen by a more basic nitrogen, with functional consequences down the line for H-bonding tautomer populations in nucleic acids.

The two amination steps, IMP to AMP vs. XMP to GMP, employ parallel but distinct strategies. As shown in Figure 9.4B, the AMP synthase traps out the minor isoamide anion as the phosphate ester, awaiting attack of the amine nucleophile. Surprisingly, it is aspartate in a reprise of the aspartate addition–lyase elimination of fumarate just seen in IMP construction in Figure 9.4A. The GMP synthase uses ATP cleavage also to trap the rare isoamide anion but this time as the O–AMP species rather than the O–PO₃^{2–} adduct. More intriguingly, the nitrogen nucleophile in GMP synthase is nascent ammonia from *in situ* glutaminase action: two complementary strategies to achieve the same end.

To sum up the origin of the five nitrogen atoms in AMP vs. GMP (and so in ATP and GTP), in AMP, two come from NH₃ *via* glutaminases, one from glycine and two from aspartate. In GMP, it is three nitrogens out of five that arise from *in situ* released NH₃, one from glycine and one from aspartate.

One final connection between the amide, amidine, urea and guanidine groups in the purine heterocycles is evidenced by the GMP reductase reaction. As shown in Figure 9.4C, the GMP reductase is also a deaminase. The deamination follows after hydride transfer into the guanidine, an unusual chemical step, but one that generates a guanine-derived C₂ carbon that is sp³ hybridized with three attached nitrogen atoms. This carbon with three heteroatoms is unstable and has a ready path for decomposition. As shown, elimination of the exocyclic amine as NH₃ occurs. The enzyme is an unusual deaminase: the guanidine group has been reductively deaminated back to the amidine of IMP, bypassing the XMP that was an intermediate

in the forward two-step conversion of IMP to GMP.

9.4 Cyclic Amines, Amides and Other Nitrogen Heterocycles

Figure 9.5 depicts four types of nitrogen heterocyclic metabolites found in different organisms. For example, cyclic amines range from three-member to six-member rings (**Figure 9.5A**). The smallest, three-member aziridine is seen in aziridine carboxylate-bearing peptides, while the four member azetidine carboxylate is a known, albeit rare, amino acid metabolite. **8** Proline and pipecolate make up the common five- and six-member cyclic amine scaffolds examined in previous chapters. The cyclic scaffold imposes conformational constraints but the amine nitrogens still function as nucleophiles. The three-member aziridine ring is sufficiently strained (like the corresponding epoxide oxygen framework) that it imparts electrophilic properties to the adjacent carbon as in the streptomycete metabolite mitomycin family **9** that act as bioalkylating agents by way of the aziridine ring opening.

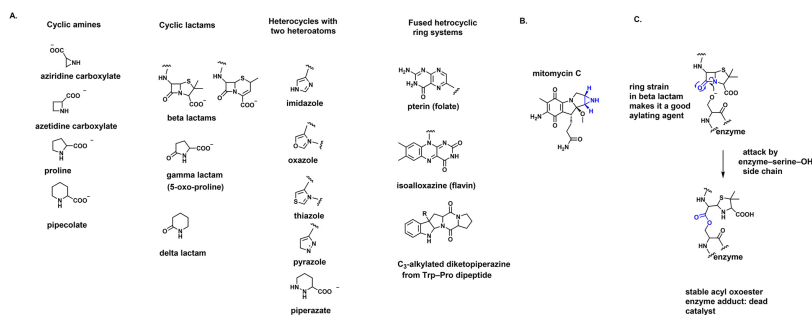


Figure 9.5 (A) Common cyclic metabolites: 3–6-membered cyclic amines; 4–6-membered lactams 5- and 6-membered heterocycles with two heteroatoms, fused heterocyclic ring systems. (B) The molecular basis of beta lactams as antibiotics is the strained four-membered lactam ring that is thermodynamically activated but kinetically stable in aqueous solutions until it binds to targeted bacteria cell wall enzymes, where it undergoes ring-opening by an enzyme active site serine side chain hydroxyl group. (C) Penicillin as a reactive acylating agent.

Among the cyclic amides (= lactams) in column 2 of **Figure 9.5A**, the β -lactam, thermodynamically strained four-membered ring, is famous as the chemical warhead in the penicillin and cyclosporin class of antibiotics (**Figure 9.1C**), the most widely prescribed antibiotic class over the past 80 years. **10–11** The nitrogen in the amide functional group is adjacent to the electrophilic carbonyl group that is the site of nucleophilic attack. In this antibiotic class, the protein targets are the serine side chains of active site residues of bacterial cell wall transpeptidases and carboxypeptidases (**Figure 9.5C**). The initial tetrahedral adduct can breakdown in the reverse or forward direction. In the forward direction the result is a covalent oxo ester to the enzyme active site serine side chain that can be stable for hours, long enough to leave the bacterium vulnerable to bursting from internal turgor pressure. In contrast, the five- and six-membered gamma and delta lactams are free of comparable ring strain and are generally unreactive cyclic amides in physiological milieus.

The urea, amidine (imine derivative of amide) and guanidine groups (imine

derivative of urea (not shown)) feature prominently as part structures in heterocycles, most notably the pyrimidine and purine scaffolds noted in the previous sections of this chapter.

Most of the heteroaromatic frameworks with significant roles in biology are five- and six-membered ring structures with one or two nitrogen atoms. As noted in the next chapter, seven of the eight B vitamins employ chemistry enabled by such nitrogen heterocycles to function as coenzymes in central metabolic transformations. In those contexts, the monocyclic, bicyclic, tricyclic and tetracyclic heteroaromatic nitrogen frameworks are the business ends of individual coenzymes.

The third column in [Figure 9.5A](#) lists some common five- and six-membered heterocyclic systems with either two nitrogens or a nitrogen and an oxygen or sulfur atom. Proceeding down the column from imidazole are three related heterocycles that show up much less often in biological systems. Replacement of one of the imidazole nitrogens by oxygen yields the oxazole ring. During assembly of oxazole frameworks in peptides and proteins, the pathway proceeds through the dihydroheteroaromatic oxazoline. Analogously, replacement of one nitrogen in imidazole by sulfur yields the thiazole ring and its biosynthetic thiazoline precursor. In contrast to the imidazole-containing free amino acid histidine, the oxazole and thiazole rings are assembled only on preformed serine, threonine or cysteine peptides. In microbial metabolism the dihydro precursor oxazoline and thiazoline and the aromatized oxazole and thiazole rings can function as ligands to ferrous atoms in nonribosomally generated bacterial peptide siderophores. [12](#)

Movement of the nitrogen at the 3-position of the imidazole ring to the 2-position turns the common imidazole heterocycle into the 1,2-pyrazine ring system. Pyrazolylalanine is a known but rare nonproteinogenic amino acid metabolite in microbial metabolism. [13](#) Analogous six membered rings are termed pyrazines. We will note a 1,4-tetrahydropyrazine ring system in the active form of vitamin B₉ in [Chapter 10](#).

The fourth column in [Figure 9.5A](#) shows three fused nitrogen heterocyclic systems along with the alkaloids shown in [Figure 9.1D](#). The bicyclic pterin ring system and the tricyclic isoalloxazine ring system are part of the B vitamin class discussed in the next chapter. The formidably complex-looking five ring system that is the last entry in column four is a simple two-step metabolite from an L-tryptophan–L-proline dipeptide and leads into the next section on indole ring chemical biology.

9.5 Tryptophan and Tryptamine as Latent C₃ Carbanion–Enamine Equivalent

9.5.1 Tryptamine: Acylation, Alkylation and Iminium Formation

The two nitrogens in tryptophan, the C₂-amino group and the indole ring N–H, participate differently as precursors to thousands of plant and microbial metabolites. One early path involves enzymatic processing by pyridoxal phosphate-dependent tryptophan decarboxylases to CO₂ and tryptamine [14](#) ([Figure 9.6A](#)). That amine undergoes two subsequent types of reactions with cellular electrophiles.

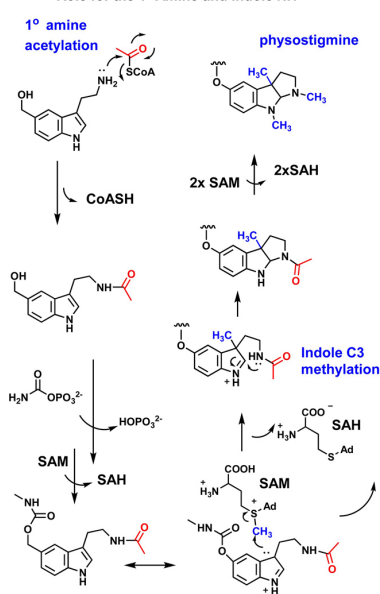
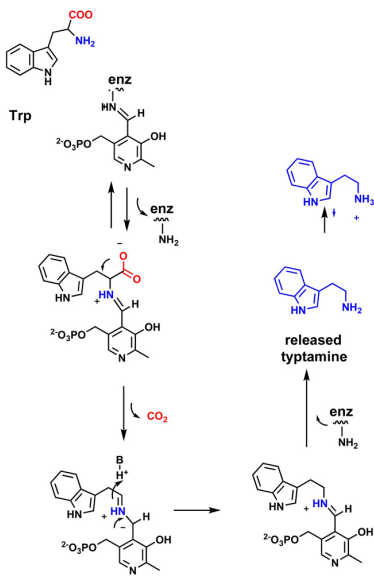


Figure 9.6(A) Pyridoxal phosphate-dependent decarboxylation of L-tryptophan to tryptamine. (B) Physostigmine formation: *N*-acetylation, C₃-methylation, *N*-deacetylation, two *N*-methylations.

9.5.2 Physostigmine Biosynthetic Strategy

One metabolic pathway features *N*-acylation, *N*-deacylation and *N*-alkylation of the primary amine and *N*-alkylation of the indole nitrogen, exemplified during biosynthesis of the anticholinesterase metabolite physostigmine. ¹⁵ Tryptophan is first hydroxylated at the *para*-ring position by a nonheme iron and tetrahydrobiopterin-dependent hydroxylase followed by decarboxylase action. As illustrated in [Figure 9.6B](#), the primary amine of 5-OH-tryptamine attacks the electrophilic acetyl group of acetyl CoA to form a classic *N*-acetylamide linkage (see Chapter 6).

A third nitrogen is introduced as an O-carbamyl group using the cellular carbamylating agent carbamyl phosphate (see Chapter 7). After carbamylation and then methylation at the phenolic oxygen, the indole nitrogen reveals its chemical capability *via* the enamine moiety at the N₁-C₃ of the pyrrole ring component.

The indole NH lone pair can participate in formation of a C₃ carbanion equivalent (Figure 9.6B). That β-carbanion carries out a large range of prototypic indole-mediated biological transformations. In this case C₃ attacks the [CH₃]⁺ equivalent of cosubstrate S-adenosyl methionine (SAM). This traps the indole scaffold as a methylated indoline that has an iminium part structure at C₂=N₁ as an electron sink in the pyrrole moiety of the indole framework. This can be captured by a nearby nitrogen nucleophile, apparently even a weak one in the guise of the original amino nitrogen, now tied up as the acetyl amide. This C₃ methylation (C-alkylation), followed by the C₂ iminium capture, generates a new N–C₂ bond as the bicyclic indole is converted into a tricyclic 5,5,6- framework, a *pyrroloindole* scaffold. This is the general strategy for hundreds of microbial pyrroloindole tricyclic metabolite scaffolds: C₃ alkylation followed by (weak) nitrogen nucleophile addition to the indole C₂ iminium carbon.

The pathway to physostigmine continues with side chain N₁-deacylation and then

two further methylations, on each of the two nitrogen atoms. In the physostigmine tricyclic structure the two nitrogen atoms from the original tryptamine framework end up as tertiary amine derivatives. The third nitrogen, introduced as a carbamyl group is a secondary *N*-methylated group.

To reprise, the two nitrogens of tryptamine, primary amine and indole NH, are each utilized for specific chemical transformations. The primary amine is first *N*-acylated, then acts as nitrogen nucleophile towards C₂ of the C₃ alkylated indole, before undergoing deacylation and behaving as nucleophile for *N*-methylation. *N*-acylation, *N*-deacylation and *N*-alkylation show a nitrogen protection–deprotection strategy along the biosynthetic route. (This competitive *N*-acylation vs. *N*-alkylation for capture of a nucleophilic amino group prefigures histone lysine side chain amine acylation–alkylation competition – Chapter 16.).

In parallel, the indole nitrogen facilitates a C₃ carbanionic *C*-methylation, acts as an electron sink in the ring closure to the tricyclic pyrroloindole skeleton, and then is *N*-methylated. Thus, there are seven transformations of those two nitrogen atoms in the nucleophilic mode, respectively. Along the way a bicyclic 6,5-indole framework has been turned into a fused 6,5,5-tricyclic pyrroloindole. The new ring is an appended tetrahydro-pyrrole – a pyrrolidine.

9.5.3 Imine to Beta Carboline Scaffold

A characteristic second mode of reactivity of tryptamine in plant and microbial metabolic pathways is for the primary amine to interact with an *aldehyde* cosubstrate and undergo initial *aldimine* formation. ² For example, in [Figure 9.7](#) acetaldehyde forms the indicated aldimine. Again, the indole NH participates as an enamine reagent at C₃ to form the indole β-carbanion for intramolecular attack on the imine carbon, creating a new C–C bond. This is well known in abiotic organic chemistry as the Mannich reaction. ¹⁶ It is the start of a number of secondary metabolite pathways. One mechanistic formulation for the fate of the initial Mannich adduct is decomposition with C–C single bond migration to produce the 6,5,6 tricyclic scaffold known as a tetrahydro-β-carboline. Oxidation to the heteroaromatic ring system is facile, yielding the characteristic carboline nucleus.

Enzymatic Mannich Condensation

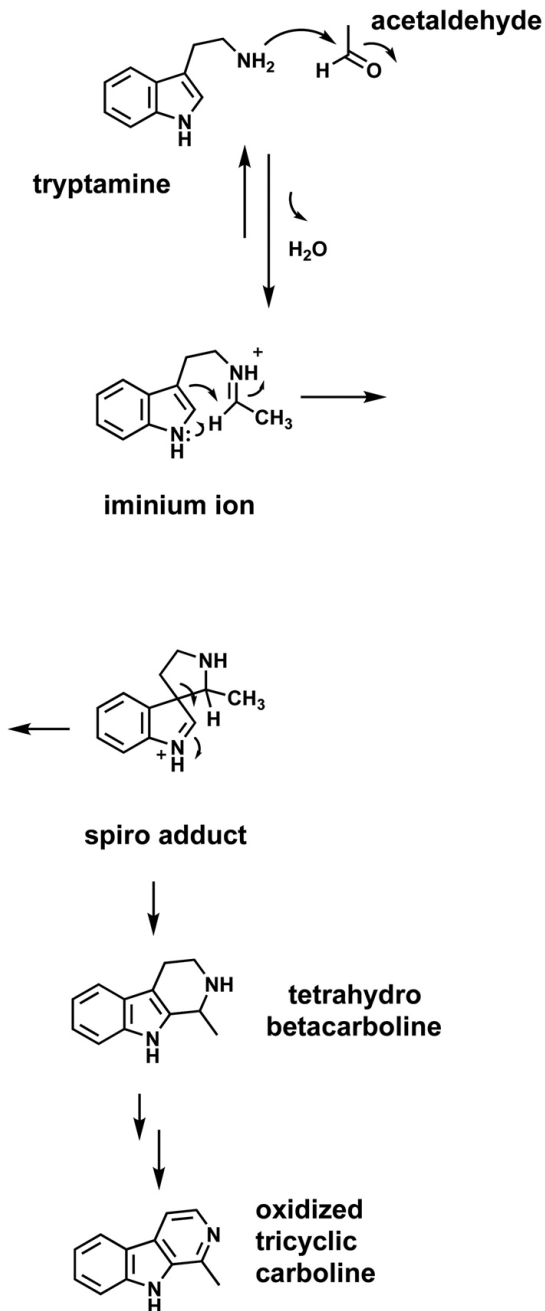


Figure 9.7 Mannich reaction of tryptamine and acetaldehyde leads to 6,5,6 tricyclic carboline scaffold.

Thus, acylation of the primary amine in physostigmine formation led to a 6,5,5

tricyclic scaffold (Figure 9.6B). Initial imine formation and Mannich reaction has led to the distinct 6,5,6 tricyclic carboline ring system (Figure 9.7). In each case the indole nitrogen as an enamine part structure has facilitated incipient indole beta carbanion formation for N–C or C–C bond formation, respectively.

9.5.4 Tryptophan: Diketopiperazine Formations and Internal Cyclizations

An additional common route to indole alkaloid scaffold biosynthesis involves initial formation of Trp–X dipeptidyl intermediates on nonribosomal peptide synthase assembly lines (Chapter 5), followed by intramolecular cyclization as the release mode 17 as shown in Figure 9.8A. Two examples are an L-Trp–L-Pro dipeptidyl–S-enzyme assembly line and an L-Trp–L-Ser dipeptidyl–S-enzyme assembly line. If the Trp-amino group acts as a kinetically competent nucleophile on the tethered thioester carbonyl, the released product will contain two amide bonds. One is the original peptide bond, between Trp–Ser or Trp–Pro. The second amide bond is newly formed between the Trp–NH₂ and the Ser or Pro activated thioester carbonyl. The released product then contains a six-membered ring with two nitrogen atoms in a 1,4- relationship. Both are in amide linkage. The ring system is called a diketopiperazine (DKP) 17 (Figure 9.8A). (Piperazine is the parent 1,4-ring system. The two carbonyls are termed ketos even though they are in fact amide carbonyl groups.)

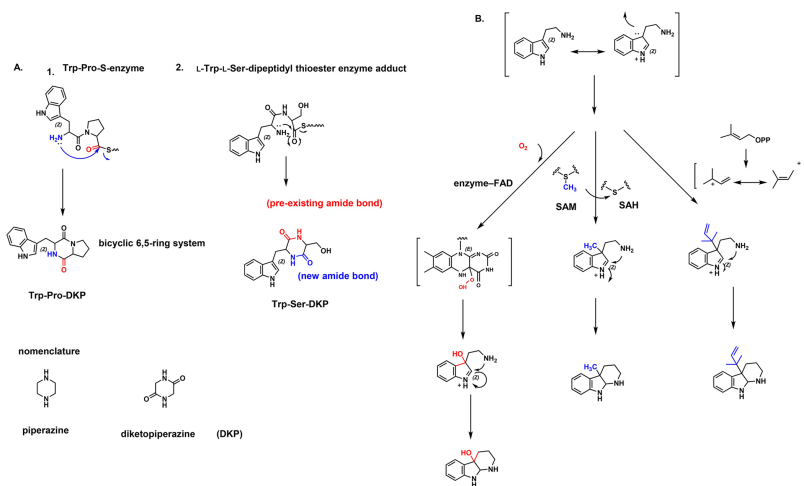


Figure 9.8(A) Tryptophan-containing dipeptidyl enzyme internal release to cyclic diketopiperazine derivatives. (B) Enzymatic capture of three cellular electrophiles by the enamine moiety of tryptamine at C₃, followed by C₃-hydroxylation and cyclization to 6,5,6 tricyclic system; comparable capture of two carbon electrophiles, the methyl group of SAM and the C₃-allylic carbenium ion from dimethylallyl diphosphate.

Diketopiperazines are common dipeptide analogs in plant and microbial metabolism. These are starting points for further elaboration, usually *via* the indole–enamine C₃ carbanion reactivity mode. The question arises as to what cellular electrophiles are available to be captured by indole acting in an enamine reactivity mode. 18 One common carbanion electrophile is the methyl group of S-adenosyl methionine as noted in the physostigmine biosynthetic pathway. Two other electrophiles are (1) carbon 1 of

prenyl diphosphates and (2) an oxygen electrophile, the $[\text{OH}]^+$ equivalent from flavin-4-a-OOH intermediates, 18 noted in the *N*-oxygenation Chapter 12 (Figure 9.8B).

Most common is capture of dimethylallyl-diphosphate, the major five-carbon alkylating metabolite in all of isoprenoid and steroid biosynthesis. Early dissociation of the $\text{C}_1\text{-OPP}$ ($\text{C}_1\text{-O}$) bond is facilitated because the resulting C_1 carbocation is delocalized into the $\text{C}_2=\text{C}_3$ double bond by resonance. It is a delocalized, stabilized *allylic cation* (Figures 9.8B and 9.9), with charge deficiency at both C_1 , the primary carbon, and C_3 , a tertiary carbon (more stable cation).

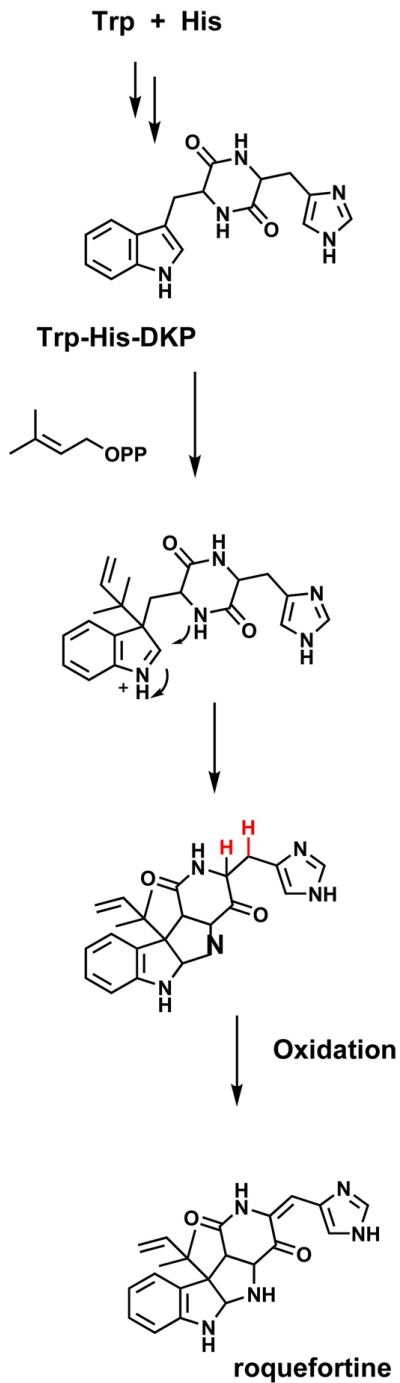


Figure 9.9 Roquefortine formation from Trp and His via prenylation of the diketopiperazine intermediate.

Capture of the five-carbon dimethylallyl carbocationic transition state by the Trp-His

diketopiperazine is directed by an enzyme in *Penicillium roquefortii* as shown in Figure 9.9. The initial adduct is the C₃-prenylated isomer with a formal C₂=N⁺ electrophilic center. As in the physostigmine example above, the weakly nucleophilic, but highly proximate, DKP amide nitrogen attacks C₂ to quench the charge on the imine. There is now a new C–C bond at C₃ and a new C–N bond at C₂ of what started as the indole framework. This is an extended contiguous 6,5,5,6 tetracyclic scaffold. This molecule is roquefortine D. 19 . (Note that the indole C₃ carbanionic equivalent captured the prenyl group at its C₃ cationic resonance contributor to give the indicated structure of roquefortine D.)

9.5.5 Alternate Regiochemistry of Indole Prenylations

C₃ prenylation of the indole ring of Trp-X-diketopiperazines is relatively common regiochemistry for biological alkylation of the indole ring in tryptophan or tryptamine frameworks. In fact, there are examples of metabolites prenylated regioselectively at C₇, C₆, C₅, C₄, C₃, C₂, and even N₁. 20 That is, all six carbons and the indole nitrogen can be alkylated selectively by enzymes using the five-carbon cationic precursor dimethylallyl diphosphate. Indeed carbon 4 is a prime site for indole alkaloid prenylation as in the lysergic acid structural series 2 (Figure 9.10A).

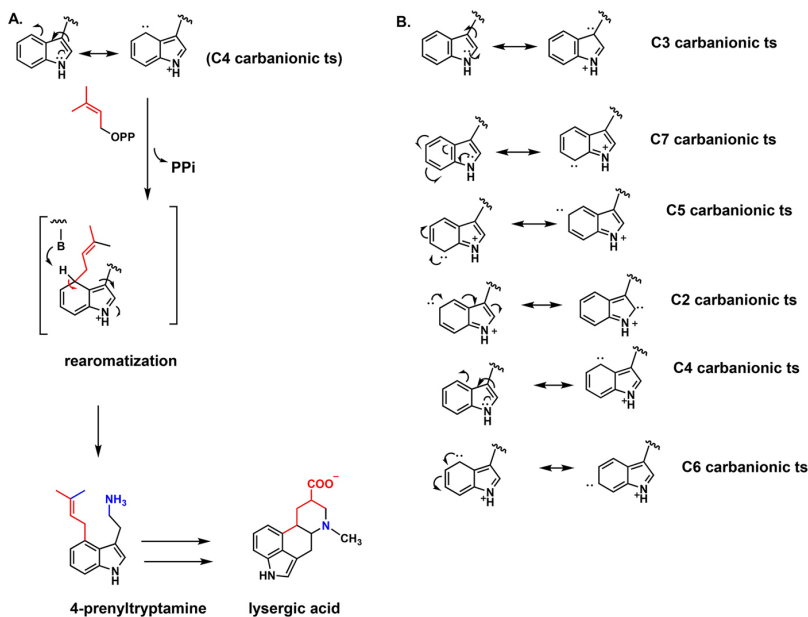


Figure 9.10 Tryptophan can behave as a carbon nucleophile at six carbon sites around the indole ring. (A) Schematic of 4-prenyltryptamine formation on the way to lysergic acid. (B) Resonance contributors to carbanion density at each of the indicated carbon sites on the indole ring (carbanionic ts = carbanionic transition state).

One can write a series of formal resonance structures (Figure 9.10B) that show how the lone pair on the indole nitrogen can be delocalized to specific carbon atoms to account for the ability of indole prenyltransferases to attach the prenyl group (either at C₁ or C₃ of the prenyl moiety in flight) to any of the indole ring positions. The 4-prenyl

indole isomer is depicted in [Figure 9.10A](#) as arising from one of those resonance contributors.

Separately, a variety of regiospecific bacterial chlorinases can make 5-chloro-, 6-chloro-, or 7-chloro-indole isomers, by delivering the equivalent of $[\text{Cl}]^+$ by way of an intermediate FAD 4a-OCl adduct. A schematic route to 7-chlorotryptophan is shown in [Figure 9.11A](#).

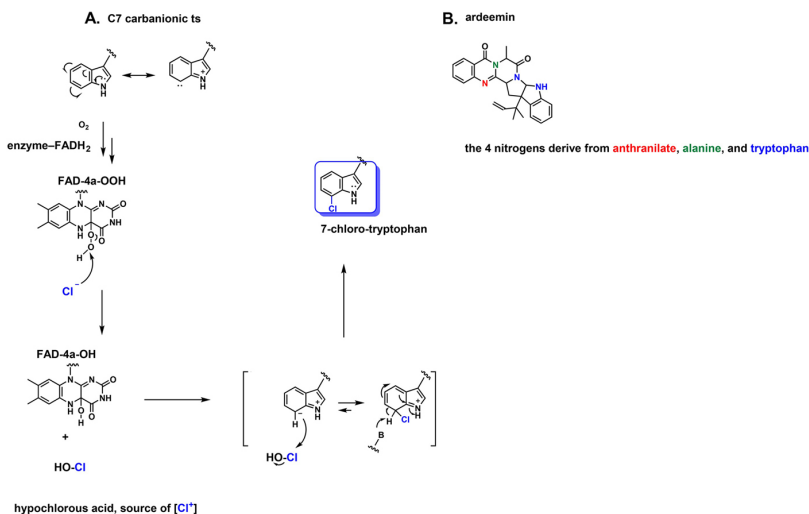


Figure 9.11 Tryptophan can behave as a carbon nucleophile at six carbon sites around the indole ring. (A) Schematic of O_2 - and FAD-requiring halogenase to generate 7-chloro-Trp. (B) Ardeemin is a fused hexacyclic scaffold with four nitrogen atoms from the three amino acid building blocks anthranilate, alanine and tryptophan.

Readers interested in accessing the literature of alkylated, acylated and oxygenated natural products in scaffolds that reach hexa-heterocyclic fused ring scaffolds, such as the fungal peptidyl alkaloid ardeemin ([Figure 9.11B](#)), might consult the reference by Walsh and Yang [2](#) for entry into that natural product biosynthetic literature.

9.6 Overview

The core chemical logic for information storage in both DNA and RNA is embedded in the amide–amidine–urea–guanidine functional groups in the nitrogen heterocyclic pyrimidine and purine building blocks of the nucleic acids. The Watson–Crick base pair framework of two H bonds between A–T pairs and three H bonds between G–C pairs represents the translation of nitrogen chemistry into bioinformation. The amide, amidine, urea and guanidine functional groups have been involved since the beginning of life on earth and probably in prebiotic assembly of nitrogen heterocycles. For example, adenine has the chemical formula $\text{C}_5\text{H}_5\text{N}_5$ and can assemble from five molecules of HCN under prebiotic conditions.

Perhaps in keeping with the presence of pyrimidine and purine bases in primitive organisms, the biosynthetic pathways are simple both in the inventory of starting substrates and the type of chemical steps in the execution of ring biosyntheses. The sources of the nitrogen atoms that eventually end up in amide, amidine, urea or

guanidine linkages are carbamyl phosphate, glycine and aspartate and the carboxamide of glutamine. Indeed, three of the five nitrogens in guanine nucleotides derive from *in situ* glutaminase liberation of nascent NH₃. In adenine nucleotides, two nitrogens come *via* the glutaminase route, two from the aspartate α -amino group.

The other thread in organic amine functional group chemical biology is iminium ion formation between amines and carbonyl groups. When one of the carbons α -to the iminium carbonyl group has a C-H bond, that hydrogen is abstractable in a low energy step as a proton to form the enamine tautomer, a stabilized form of the carbanion. The indole group in tryptophan and its decarboxylation product tryptamine are the starting points for thousands of plant- and microbial-based secondary metabolites by virtue of enamine chemical biology. Indoles with their enamine part structure at N₁-C₃ in its pyrrole ring are well known for facile generation of C₃ carbanionic transition states on the way to tricyclic pyrroloindole and higher order fused indole ring systems. Indeed, alkylation, acylation and hydroxylation at C₃ of the indole ring are common biosynthetic operations in plant and microbial secondary metabolic pathways.

We continue exploration of key nitrogen heterocycles that power many primary metabolic pathways in the next two chapters. Chapter 10 deals with nitrogen heterocycles in the coenzyme forms of the B vitamin family. Chapter 11 takes up tetrapyrrole macrocycles, sometimes called the pigments of life.

References

1. C. T. Walsh Nature loves nitrogen heterocycles, *Tetrahedron Lett.*, 2015, **56**, 3075-3081.
2. C. T. Walsh and Y. Tang, *Biosynthesis of Natural Products*, Royal Society of Chemistry Press, London UK, 2019.
3. *Watson-crick Hydrogen Bonds: Nature and Role in DNA Replication*, ed. C. F. Guerra and F. Bickelhaupt, Elsevier, 2006, ch. 4.
4. M. E. Jones Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis, *Annu. Rev. Biochem.*, 1980, **49**, 253-279.
5. S. C. Sinha and J. Smith The PRT protein family, *Curr. Opin. Struct. Biol.*, 2001, **11**, 733-739.
6. F. M. Raushel, J. Thoden and H. M. Holden The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia, *Biochemistry*, 1999, **38**, 7891-7899.
7. Y. Zhang, M. Morar and S. E. Ealick Structural biology of the purine biosynthetic pathway, *Cell. Mol. Life Sci.*, 2009, **65**, 3699-3724.
8. L. Fowler Azetidine-2-carboxylic acid a new cyclic imino acid occurring in plants, *Biochem. J.*, 1956, **64**, 323-331.
9. G. S. Kumar, R. Lipman, J. Cummings and M. Tomasz Mitomycin C-DNA adducts generated by DT diphorase. Revised mechanism of the enzymatic reductive activation of mitomycin C, *Biochemistry*, 1997, **36**, 14128-14136.
10. M. M. Miyachiro, C. Contreras-Martel and A. Dressen Penicillin binding proteins (PBPs) and bacterial cell wall elongation complexes, *Subcell. Biochem.*, 2019, **93**, 273-289.
11. C. T. Walsh and T. Wenczewicz, *Antibiotics: Challenges, Opportunities*, ASM Press, Washington DC, 2016, p. 500p.
12. J. H. Crosa and C. T. Walsh The genetics and assembly line enzymology of siderophore biosynthesis, *Microbiol. Mol. Biol. Rev.*, 2002, **66**, 223-249.
13. P. M. Dunnill and L. Fowler The biosynthesis of beta-pyrazolyl-1-alanine, *J. Exp. Bot.*, 1963, **14**, 237-248.
14. P. K. Mehta and P. Christen The molecular evolution of pyridoxal-phosphate-dependent enzymes, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1999, **74**, 129-184.

15. J. Liu, T. Ng, Z. Rui, O. Ad and W. Zhang Unusual acetylation-dependent cascade in the biosynthesis of the pyrroloindole drug physostigmine, *Angew. Chem., Int. Ed.*, 2014, **53**, 136-139.
16. F. F. Blicke The Mannich Reaction, *Org. React.*, 2011, **1**, 303-341.
17. J. A. Baccile, H. Henry Le, B. T. Pfannestiel, J. W. Bok, C. Gomez, E. Brnadenburger, D. Hoffmeister, N. P. Keller and F. C. Schroeder Diketopiperazine formation in fungi requires dedicated cyclization and thiolation domains, *Angew. Chem., Int. Ed.*, 2019, **58**, 14589-14593.
18. C. T. Walsh, S. Haynes, B. D. Ames, X. Gao and Y. Tang Short pathways to complexity generation: fungal peptidyl alkaloid multicyclic scaffolds from anthranilate building blocks, *ACS Chem. Biol.*, 2013, **8**, 1366-1382.
19. H. Ali, M. Ries, J. G. Nijland, P. P. Lankhorst, T. Hankemeier, R. A. L. Bovenberg, R. J. Vreken and A. J. M. Driessen A branched biosynthetic pathway is involved in the production of roquefortine and related compounds in *penicillium chrysogenum*, *PLoS One*, 2013, **8**(6), e65328.
20. C. T. Walsh Biological matching of chemical reactivity: pairing indole nucleophilicity with electrophilic isoprenoids, *ACS Chem. Biol.*, 2014, **9**, 2718-2728.

B Vitamin Nitrogen Heterocycles

10.1 Seven Human B Vitamins Contain Functional Nitrogen Heterocycles

Human vitamins are defined as molecules required exogenously as dietary micronutrients to sustain health in acute and chronic conditions. [1](#) If the pyrimidine and purine heterocycles in RNA and DNA are one apotheosis of the utility of nitrogen heterocycle chemistry to enable life and information transfer, then the seven B vitamin nitrogen heterocycles are the radiation of that nitrogen chemistry to almost every major metabolic pathway in humans. [2](#)

All 13 known human vitamins were discovered in the first half of the twentieth century based on nutritional studies that associated human disease syndromes with a nutritional deficiency that could be cured by adding the missing molecule back to the diet. Before molecules were fully purified and vitamin structures established, there was an initial (incorrect) hypothesis that all vitamins were amines, *e.g.* “vital amines” which gave rise to nomenclature for the whole diverse nutritional class.

In general, as each new vitamin was isolated and characterized it received a capital letter designation: vitamin A, vitamin B, vitamin C, vitamin D, *etc.* Another categorization of the full human vitamin list is that four vitamins, A, D, E and K, are lipophilic and so lipid soluble. The other nine are hydrophilic and water soluble. As the physiological mechanisms of each human vitamin were deciphered, it became clear that eight of the nine water soluble vitamins functioned as coenzymes for specific classes of enzymes, providing the key chemical functionalities for the enzyme–coenzyme complexes to catalyze certain types of metabolic transformations. All those eight vitamins acting as coenzymes were grouped together as B vitamins. The remaining water-soluble vitamin, vitamin C, escaped this lumping together, but it can actually function as a cofactor, for example in proline hydroxylases involved in collagen fiber maturation.

One might expect that the eight B vitamins would be neatly numbered from B₁ to B₈, but that is not the case because during the half century of active vitamin hunting, some candidates fell by the wayside. In particular vitamins B₄, B₈, B₁₀ and B₁₁ did not stand the test of time, either as an authentic vitamin or sometimes because the substance was actually an endogenous metabolite. Thus, the B vitamin list for the past 70 years has been B₁ (thiamin), B₂ (riboflavin), B₃ (niacin = nicotinamide), B₅ (pantothenic acid), B₆ (pyridoxine), B₇ (biotin), B₉ (folate) and B₁₂ (cobalamin). [Figure 10.1](#) lists the structures of all eight B vitamins, while [Figure 10.2](#) summarizes human diseases associated with specific vitamin nutritional deficiencies. Since the 1950s each of the B vitamins can be supplied as pure dietary supplements. However,

all of them are converted *in vivo* to coenzymatically active forms, as will be discussed in subsequent sections.

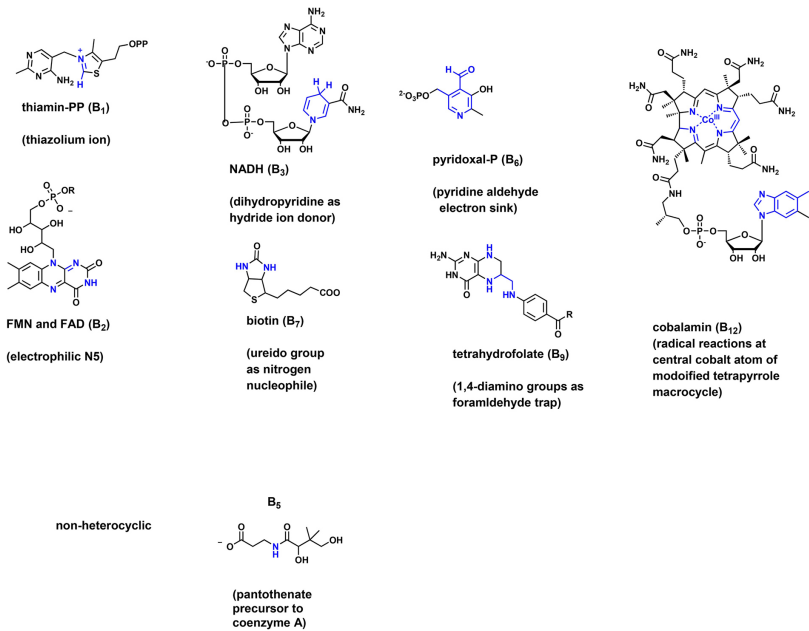


Figure 10.1 Nitrogen heterocycles at the active core of the coenzyme forms of seven out of eight B vitamins.

Disease	Deficiency
beri beri	B₁ (thiamin)
stomatitis	B₂ (riboflavin)
pellagra	B₃ (niacin)
anemia, peripheral neuropathy	B₆ (pyridoxine)
conjunctivitis	B₇ (biotin)
megablastic anema	B₉ (folic acid)
pernicious anemia	B₁₂ (cobalamin)

Figure 10.2 Human diseases arising from nutritional deficiency of specific B vitamins.

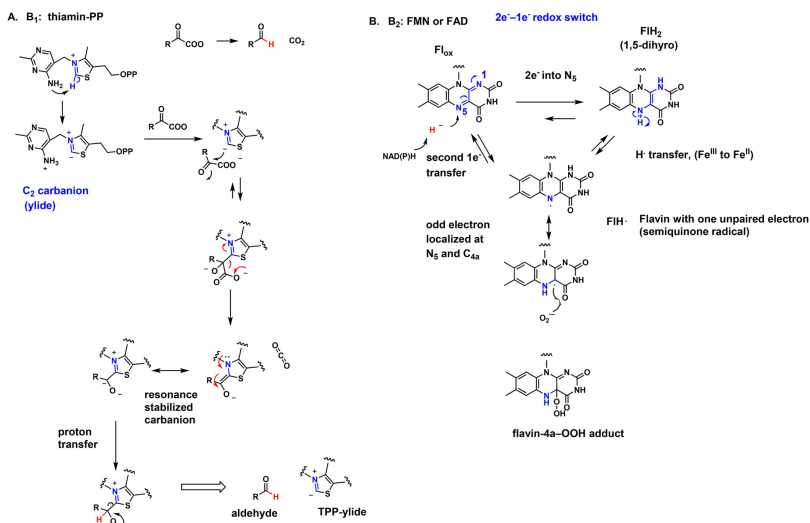
Inspection of the structures of the eight B vitamins, as a prelude to mechanism shows that seven of the eight have heterocyclic nitrogen atoms. Pantothenic acid, vitamin B₅, the metabolic precursor to coenzyme A, has one nitrogen in the form of an acyclic amide between β -alanine and the pantoate scaffold. It is the sole B vitamin where a heterocyclic nitrogen atom is not at the center of catalysis.

Over and above the nonreacting AMP portions thought to be binding entities for conferring high affinity to target enzymes, there are seven distinct nitrogen heterocycles that function as the catalytic moieties, each of which is specialized for unique chemistry in association with partner apoprotein constituents of enzyme-coenzyme complexes.

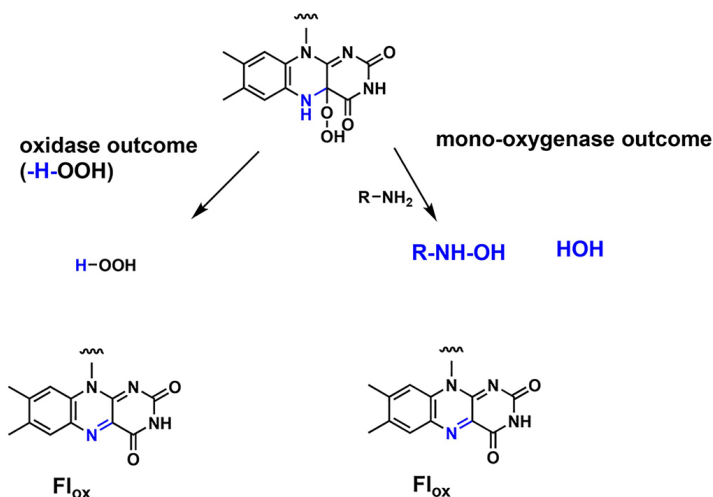
The seven heterocycles illustrate the breadth of both nature's ingenuity at heterocycle assembly and the chemical versatility of those nitrogen heterocycles 2:3 (Figure 10.1). They include the permanently charged thiazolium ion in B₁ and the quaternized pyridinium ion in the nicotinamide business end of the B₃ coenzyme forms. The pyridine aldehyde is the workhorse group of the pyridoxal phosphate B₆ coenzyme while the ureido ring of biotin (B₇) is the carrier of the one-carbon fragment at the oxidation state of CO₂. In complementarity, the *para*-aminobenzoyl-pterin scaffold of reduced folate coenzyme (B₉) is the cellular carrier of one-carbon fragments at the oxidation states of methanol, formaldehyde and formic acid. Thus, B₇ and B₉ can ferry all four of the one-carbon oxidation states between donor and acceptor substrates.

The tricyclic isoalloxazine ring system of the vitamin B₂ flavin-based coenzymes

equips N^5 as an electrophilic nitrogen in the oxidized state and a nucleophilic nitrogen in the two-electron reduced state (Figures 10.3B and 10.4). Furthermore, the tricyclic framework acts as a two-electron–one-electron switch, interfacing between obligate two-electron donors [NAD(P)H] and obligate one-electron acceptors (Fe^{III} and O_2). For vitamin B_{12} we will note how the tetrapyrrole macrocycle has been customized to fit cobalt ions as the equatorial ligand to the four pyrrolic nitrogens and to control the electronic microenvironment around cobalt.



Figures 10.3 Mechanisms of action for coenzyme forms of vitamins B_1 and B_2 . (A) Thiazolium ring of thiamin-PP as source of C_2 carbanion to add to substrate carbonyl groups; (B) tricyclic isoalloxazine ring system of flavin coenzymes as step-down redox transformer between two-electron donors and one-electron receivers.



Figures 10.4 Mechanisms of action for coenzyme forms of vitamins B₁ and B₂. Partitioning of Flavin-4a-OOH between oxidase and oxygenase modes.

10.2 The Quaternized Thiazolium Ring in Thiamin-PP the Coenzyme Form of Vitamin B₁

Thiamin was the first vitamin to be isolated, structurally characterized and identified as the deficient molecule in a disease, beri beri, arising from its deficiency in polished white rice that was the staple food in the diet of certain East Asian populations. ⁴ The vitamin is in the husk layer, present in brown rice but stripped away in polished white rice. The coenzyme form is only modestly different from the vitamin, with enzymatic pyrophosphorylation of the thiamin alcoholic oxygen. The business end of the vitamin, the pyrimidine ring connected by a methylene bridge to the thiazolium ion permanently quaternized on nitrogen, is the same in both vitamin and coenzymatically active form. The C₂-H bond in that *cationic thiazolium ring* is labile and the hydrogen exchanges readily with solvent deuterium, indicative of a carbanion intermediate (Figure 10.3A). ⁵ In the active site of its partner enzymes, the C₂-H bond is labilized at rates >300 per second and the carbanion is on the pathway. ⁶

The key to the thiazolium carbanion kinetic accessibility is that the C₂ carbanion is adjacent to the cation at N₁, generating a charge separated ylide system (phosphonium ylides are intermediates in the famous abiotic Wittig reactions). The electrostatically paired positive and negative charges stabilize the C₂-carbanion form. That thiazolium C₂ carbanion acts as carbon nucleophile towards the ketone carbonyl of bound cosubstrates. Figure 10.3A depicts attack of that carbanion on the carbonyl carbon of an α -keto acid metabolite. The resulting adduct now has an iminium ion electron sink β to the substrate carboxylate. Decarboxylation from that adduct is a low energy pathway because the carbanion that would arise is stabilized by the indicated delocalization of negative charge between N₁ of the thiazole ring and the alcoholic

carbon of the adduct. Protonation at that carbanion reveals that the structure is the tetrahedral addition product of an aldehyde with the thiazolium C₂ carbanion. Collapse of that adduct gives the aldehyde product (along with the CO₂ that departed earlier in the reaction sequence) and regenerates the thiamine thiazolium C₂ carbanion for the next catalytic cycle.

In sum, the quaternization of the thiazolium ring nitrogen in thiamin enables low energy C₂-carbanion formation to become kinetically accessible. Among the virtues of coenzyme B₁ is its ability to provide a low energy pathway to decarboxylation of metabolically central α -keto acids such as pyruvate and 2-oxoglutarate. There are no other straightforward, low energy chemical mechanisms for α -keto acid decarboxylations in biology. Note that absent the positive charge on N₁, the C₂-H in nonquaternized thiazole rings is not notably acidic. The genius in the invention of vitamin B₁ was the coupling of the pyrimidine and thiazolium heterocycles in one scaffold **2** to generate the permanent charge on the thiazolium nitrogen.

10.3 The Tricyclic Ring System of Flavin (B₂) Coenzymes: Nature's Step-down Redox Transformer

Riboflavin, named for its ribityl side chain and its brilliant yellow color (latin: *flavius* = yellow) was first isolated from milk. The tricyclic business end of riboflavin is unaltered by enzymatic conversion to its two active coenzyme forms, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), by enzymatic phosphorylation and then adenylation sequentially.

Parsing the tricyclic isoalloxazine ring into its three components reveals a dimethylbenzene ring, a central pyrazine ring and a pyrimidine ring. However, the reactivity is more than the sum of parts. The central N₅ of oxidized flavins (FMN, FAD) is *electrophilic*, a rare situation for nitrogen and a key feature built into the quite byzantine biosynthetic assembly of the tricyclic scaffold by microbes and plants. **7** The relevant two-electron redox part structure includes the N₅-C_{4a}-C_{1a}-N₁ atoms, colored blue in [Figure 10.3B](#). N₅ is thus a port of entry for electrons into oxidized flavins. Most commonly, in dozens to hundreds of flavoenzymes, that packet of two electrons comes as a hydride ion from NAD(P)H (see next section). That passage of H⁻ results in the two-electron reduced FlH₂ system. We show N₁ in FlH₂ structure as the protonated N₁H (to emphasize the FlH₂ oxidation state), but actually its pK_a is 7, such that the electron rich N₁ anion is equally available at neutral pH in cells.

Now, it is in the FlH₂ oxidation state that the flavin tricyclic ring system displays its almost unique virtuosity for an organic nitrogen metabolite. It could deliver a hydride back to a cosubstrate (such as NAD⁺ or to an imino acid), but alternatively it can give up *one electron* at a time, to obligate one-electron transfer agents. Whereas NAD(P)H is an obligate two-electron transfer metabolite, there are two obligate one-electron acceptors of tremendous consequence to aerobes. One category includes ferric ions (Fe^{III}) in their distinct microenvironments, including heme groups, Fe-S clusters and nonheme mononuclear and dinuclear iron centers (see Chapter 12). The second category is molecular oxygen. As a ground state triplet, O₂ is extremely sluggish in reaction with spin-paired organic molecules that react by two-electron pathways (e.g. carbenium ions and carbanions). This is the reason we can exist in an atmosphere of 21% O₂.

One-electron transfer from FlH₂ to an Fe^{III} site in an enzyme generates one-electron reduced Fe^{II} and the one-electron oxidized flavin semiquinone radical [FlH[•]]. The flavin semiquinone radical is accessible kinetically and thermodynamically because the

unpaired electron is delocalized and stabilized through the tricyclic isoalloxazine framework. [Figure 10.3B](#) shows the result of transfer of an electron and proton (perhaps as $[H]$) from the N_5-H position of FLH_2 along with one resonance contributor that has the unpaired electron at the bridgehead carbon C_{4a} .

If O_2 , rather than an Fe^{III} atom, had been the recipient of the first electron transfer from FLH_2 , the product would be one-electron reduced superoxide anion radical (pK_a of protonated superoxide = 4.8). The two radicals, $O_2^{\cdot -}$ and $FLH^{\cdot}$ then recombine in a second step. Radical recombination forms the $C_{4a}-OOH$ bond and constitutes a second electron transfer out of the flavin nucleus as the $C_{4a}-O-H$ proximal bond forms.

The covalent C_{4a} -hydroperoxide adduct is a key intermediate in flavin-oxygen chemical biology and is partitioned forward in two ways by two enzyme classes. The flavoprotein oxidases ([Figure 10.4](#)) use the N_5-H , now basic, as an internal nucleophile to expel the neighboring bridgehead $-OOH$ with protonation to yield H_2O_2 and oxidized flavin. This is the stoichiometry of amino acid oxidases noted back in Chapter 6. As the amine group is oxidized by two electrons to the imine (prior to hydrolysis to ketone and NH_3), O_2 is reduced by a net two electrons to hydrogen peroxide. The mechanistic pathway is most definitely *not* a single two-electron reduction of O_2 in one step.

The alternative fate for the $Fl-4a-OOH$ occurs in almost all [8](#) the flavoenzyme oxygenases such as the N -oxygenases of Chapter 14. A substrate nucleophile, such as the primary amine, or the tertiary amine in FMO catalyses (Chapter 12), is proposed to attack the distal oxygen of the peroxy-flavin adduct, transferring it as an $[OH]^+$ equivalent, leaving the electron pair that had been in the $O-O$ peroxy bond to be protonated in the resultant $Fl-4a-OH$ adduct. That hydroxy adduct breaks down intramolecularly with loss of the proton from the N_5-H center to produce a water molecule and reveals oxidized flavin.

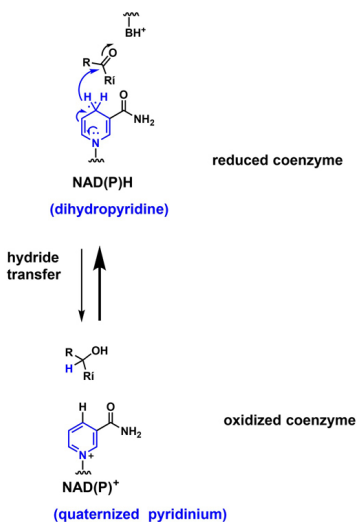
Other variants of flavoenzymes exist that are beyond the scope of this discussion. [8](#) One can summarize that oxidized flavin offers N_5 as an electrophile while in reduced flavins N_5 is both a two-electron nucleophile (*e.g.* in the intramolecular decomposition of the $4a-OOH$ -adduct) or a one electron donor of an electron (to iron, to O_2) on the way to the flavin semiquinone radical. Ribfolavin is the most versatile of the redox active vitamins and perfectly evolved as a two-electron-one-electron redox intermediary.

10.4 Dihyronicotinamides in Coenzyme Forms of Vitamin B_3 : Obligate Hydride Transfer Agents

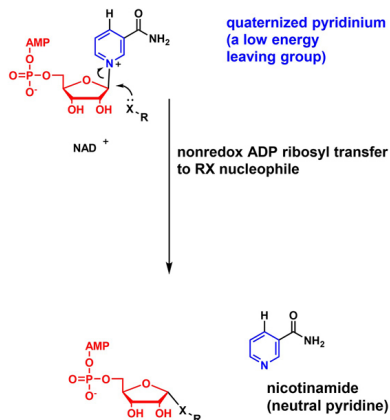
Vitamin B_3 has the chemical structure nicotinic acid, also known as niacin ([Figure 10.5](#)). It requires three enzymatic steps after human ingestion to be processed to the active coenzyme forms ([Figure 10.5B](#)). [9](#) Enzymatic phosphoribosylation by 5-phosphoribose-1-pyrophosphate (PRPP) generates nicotinic acid mononucleotide (NAMN). NAMN is then the substrate for nucleotidyl (specifically adenylyl) transfer to one of the phosphate oxygens of NAMN to produce nicotinic acid adenine dinucleotide ($NAAD^+$), an amination step away from the nicotinamide adenine dinucleotide coenzyme form NAD^+ (that amination is shown in [Figure 6.3](#), featuring the ATP requirement to make an $NAAD$ acyl phosphate intermediate captured by nascent NH_3).

A. B₃: NADH/NAD⁺ NADPH/NADP⁺

1. redox role: hydride transfer



2. ADP ribosyl transfer from NAD⁺



B.

vitamin B₃ (nicotinic acid)

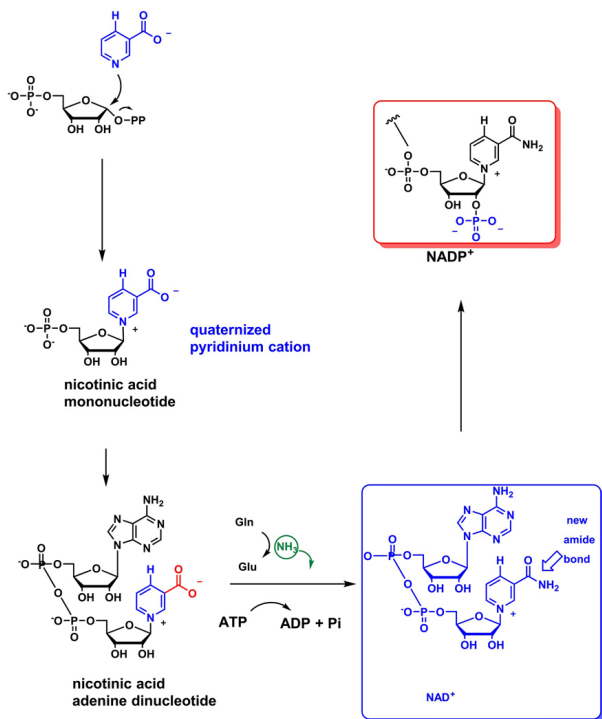


Figure 10.5 (A) Two distinct roles for oxidized nicotinamide coenzymes: redox role as hydride donor to electrophilic substrates vs. non redox role for ADP ribosyl transfers to substrate nucleophiles. (B) Steps in nicotinamide coenzyme (NAD⁺) biosynthesis from vitamin B₃ highlighting *N*-ribosyl and side chain amide bond

There are two nitrogen-relevant points to note. The first enzyme-mediated phosphoribosylation step involves attack of the pyridine ring nitrogen (not a strong base or nucleophile) on C₁ of the ribose ring of PRPP to form the pyridine N–C₁ glycoside bond. That event has quaternized the pyridine as an *N*-pyridinium ring system and is crucial for the reactivity of NAD⁺ as a hydride transfer agent, as will be noted later. The last enzyme is an ATP-dependent amide synthetase. Its chemical logic harks back to glutaminase domains in Chapter 6. To activate the resonance-stabilized carboxylate anion of the pyridinium carboxylate moiety requires first that it function as a nucleophile towards P_γ of cosubstrate Mg–ATP. The resultant pyridinium acyl phosphate is attacked by nascent NH₃ to create the neutral amide bond in the nicotinamide portion of NAD⁺. The nascent NH₃ arose from controlled glutaminase activity (Chapter 6), as usual, in a distinct domain of the NAD⁺ synthetase.

The second form of the nicotinamide coenzymes undergoes one more enzymatic step. The 2'-OH of the ribose ring attached to the quaternized nicotinamide of NAD⁺ acts as nucleophile towards P_γ of ATP to yield the corresponding 2'-phosphate known as nicotinamide dinucleotide phosphate (NADP⁺) (Figure 10.5B). NAD⁺ and NADP⁺ constitute the oxidized form of the two nicotinamide coenzymes. They are the minor oxidation state *in vivo*, where up to 99% of NADP⁺ is kept in the reduced dihydronicotinamide form, NADPH. ¹⁰ The dihydro form, NADH, also tends to predominate but the oxidized NAD⁺ is a key hydride acceptor in glycolytic metabolism.

Examination of the interconversion of reduced NAD(P)H and the two-electron oxidized NAD(P)⁺ highlights the key chemical effect of the N₁–C₁' bond attaching the nicotinamide (pyridine) ring to the ribose carbon. Whereas the oxidized NAD(P)⁺ has a permanent positive charge on the pyridine nitrogen (pyridinium), the reduced NAD(P)H is neutral because that ring is a dihydropyridine and the lone pair on the dihydropyridine nitrogen is available (Figure 10.5A). The classic role of NAD(P)H is as a hydride ion donor ¹¹ to carbonyl group cosubstrates. In the other direction, the NAD(P)⁺ oxidized form is a hydride acceptor and H[–] transfer to C₄ of the pyridinium ring results in quenching of the positive charge on N₁.

The fact that a *dihydroheteroaromatic ring* system, the dihydropyridine of NAD(P)H, is stable under physiological conditions is both remarkable and a precondition to most of the redox chemical logic of metabolism. As a dihydro-heteroaromatic ring system NAD(P)H has a strong thermodynamic driving force to give up two electrons and a hydrogen and aromatize. Thus, NAD(P)H is among the most powerful reductants in organismal chemistry. One could imagine either one-electron or two-electron transfer mechanisms for the oxidation, but the one-electron reduction potential is so far outside the biological norms that the pyridinyl radical of nicotinamide coenzymes is not relevant to biology. Its energy of formation is too high. NAD(P)H–NAD(P)⁺ is *restricted to two-electron steps only*. Furthermore, decades of experiment have indicated the pair of electrons in flight travel with the hydrogen as a hydride ion (Figure 10.5A).

The oxidation of NAD(P)H is therefore formulated as using the lone pair on N₁ of the dihydropyridine to flow into the ring to begin the aromatization process as one of the two prochiral hydrogens across the ring at C₄ is ejected as a hydride ion to a carbonyl group of a cosubstrate. The gain in resonance energy on aromatization drives the hydride ejection. In the other direction, a hydride transfer from an amine or alcoholic carbon undergoing enzymatic oxidation is directed stereospecifically to C₄ and the electrons flow to N₁ to quench the charge on that pyridinium nitrogen, converting the quaternized cation of the heteroaromatic pyridinium to the neutral, basic nitrogen of the reduced dihydropyridine ring. ¹²

Finally, the reduced, thermodynamically activated forms of NADH and NADPH are

sufficiently stable kinetically under physiological conditions that they serve as the predominant diffusible reducing currency in cells. ⁹ Because they are two-electron only redox agents and downstream cytochromes in respiratory electron transport chains are one-electron only $\text{Fe}^{\text{III}}\text{--Fe}^{\text{II}}$ transfer species, nature needed to invent two-electron to one-electron redox step-down transfer reagents to connect NAD(P)H to iron proteins. That is the role of the vitamin B_2 -based coenzymes noted above.

NAD^+ also serves in nonredox roles in cells, for ADP ribosyl group transfer to cellular nucleophiles [including DNA, protein side chains (see Chapter 16), water and even in intramolecular cyclization]. ¹³⁻¹⁴ The chemical basis for this reactivity derives again from that quaternized pyridinium ring and that bond between N_1 to C_1' of the ribose unit in the NMN portion of the coenzyme framework. Attack of cellular nucleophiles on the C_1' bond releases the free nicotinamide ring as the ADP ribosyl group is transferred (Figure 10.5A) to the cosubstrate nucleophile undergoing ADP ribosylation.

10.5 Pyridoxal Phosphate (B_6 Coenzyme): a Pyridine for all Amino Acid Transformations

The ingested form of pyridoxine, vitamin B_6 , undergoes two *in vivo* enzyme-catalyzed activation steps to convert the neutral pyridine alcohol into the functional cofactor. One step is kinase-mediated phosphorylation to the side chain alcohol phosphate as a tetrahedral anionic binding handle for partner enzymes. ¹⁵ The other, more crucial functionally, is oxidation by a flavin dependent oxidase (see Figure 10.3B) of the alcohol to the aldehyde that has two major consequences. One is that the resultant aldehyde group is a novel functional group addition to the inventory of side chains available to protein catalysts. The second consequence is that it brings the pyridine ring into conjugation with the aldehyde group, allowing the pyridine to act as an extended electron sink in subsequent enzymatic reactions (Figure 10.6).

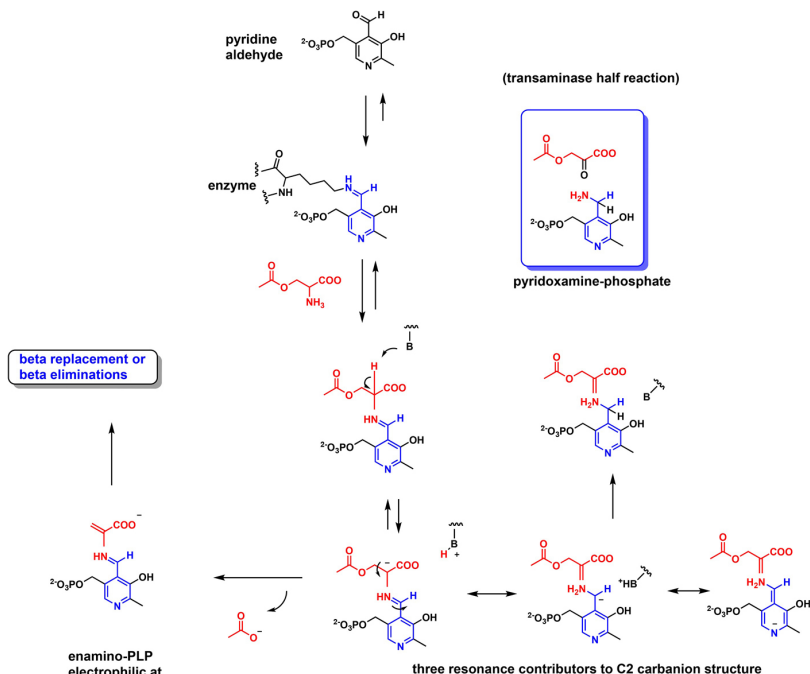


Figure 10.6 Pyridoxal phosphate, the coenzyme form of vitamin B₆, forms iminium ions with many amino acids and then serves as an electron sink to stabilize substrate–PLP carbanionic transition states. The delocalized carbanionic=PLP adduct with *O*-acetyl-*L*-serine can partition either to transamination and release of the *O*-acetyl keto acid, or undergo acetate group elimination to the enamine acid=PLP adduct on the way to either a β -replacement or β -elimination outcome.

A comparison of the key reaction centers in the coenzyme forms of B₃ vs. B₆ reveals that pyridine rings are the business ends of each coenzyme. In contrast to the pyridinium–dihydropyridine cosubstrate stoichiometry of two-electron changes in B₃, the pyridine aldehyde of B₆ stays tightly bound to its apo-enzyme partners in aldimine linkage to the N₆–NH₂ of a lysine residue. PLP *does not appear* in reaction stoichiometries (Figure 10.6). The pyridine ring in the B₆ pyridoxal coenzyme does undergo transient electron storage, especially noticeable in transaminase half reactions (Chapter 6), but is always returned to the starting heteroaromatic aldehyde state at the end of every catalytic cycle.

The pyridoxal phosphate coenzyme form is synonymous with amino acid metabolism. 16 More than 100 enzyme-mediated transformations of amino acids start out with condensation of the NH₂ group of a substrate amino acid with the resting PLP=side chain lysine iminium adduct in the enzyme's active site. The resultant amino acid=PLP aldimine adduct then engages in many variants of substrate framework carbanion chemistry directed by the orientation of the four substituents on C₂ of the amino acid moiety.

Removal of the C₂–H proton by an active site base yields the accessible, delocalized, stabilized amino acid–PLP carbanion. Electron density can be stored transiently into the pyridine heterocyclic ring, even to the point of accumulation of the *para*-quinoid resonance contributor in some enzyme active sites. The full scope of PLP–enzyme chemistry can be consulted elsewhere. The conclusion is that the conjugated pyridine

One can think of PLP as a redox coenzyme, albeit totally distinct from the pyridinium ion–dihydropyridine redox manifold of the nicotinamide coenzymes. The redox property of amino acid=PLP aldimines is particular exemplified by the two half reactions of transaminases where the PLP coenzyme is alternately reduced by two electrons to the pyridoxamine phosphate and then reoxidized in the second half reaction (please refer back to Chapter 6, Section 6.3). This redox strategy undergirds the equilibration of -NH_2 groups between different ketoacid scaffolds that accounts for the diaspora of amino group transfers from glutamate and glutamine to the other common amino acid metabolites.

When the coenzyme forms of vitamin B₇ (biotin) and B₉ (folic acid) are compared with each other they bear no discernable structural resemblance 1:2 (Figure 10.7A and B). Biotin, first isolated as a factor required for avoidance of hair loss, has two five-ring heterocycles fused, one a thiane, the other bearing the two nitrogens in a *ureido* functional group. In contrast, folate, first identified as a molecule that prevented macrocytic anemia and later shown to be essential in preventing spina bifida in developing infants, looks much closer to flavin (Figure 10.3B), with no hint of either biotin ring framework.

Figure 10.7 Functions of the coenzyme forms of vitamins B₇ (biotin) (A) and B₉ (tetrahydrofolate) (B) as intermediate carriers of one-carbon fragments at the oxidation state of CO₂ (biotin) or of HCOO⁻, HC₂=O, or CH₃-OH (tetrahydrofolate).

However, when viewed with a wider functional lens of cell physiology, one can note that biotin serves as a reversible carrier of one-carbon units at the oxidation state of CO₂ (as hydrated HCO₃⁻). In a structurally and logically distinct but functionally complementary role, the tetrahydro form of folate coenzymes carry around one-carbon units at the three different oxidation states-of methanol, formaldehyde and formate (Figure 10.8A). Between the two coenzymes, the only one-carbon fragment not carried in a latent form is methane itself. While methane is a bacterial metabolite (including in gastrointestinal bacteria), it has no known role in eukaryotic metabolism.

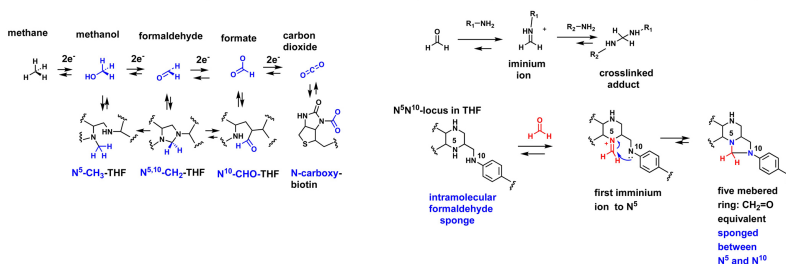


Figure 10.8(A) The five one-carbon oxidation states between methane and CO₂; (B) N⁵ and N¹⁰ of H₄-folate work in tandem as a “sponge” for promiscuously reactive formaldehyde, carrying it around as a latent unreactive, cyclic geminal diamine group (–N–CH₂–N–).

We take up the biotin and folate logic separately below but note their complementary roles as latent carriers of one-carbon units. In both coenzymes a ring nitrogen initiates attack on an electrophilic one-carbon fragment as the point of entry. The surprise is that in biotin it is a poorly nucleophilic ureido nitrogen. In tetrahydrofolate the itinerant one-carbon unit is nestled between two basic amines in a strain-free five-member ring (Figure 10.8A).

10.6.1 Biotin

Biotin becomes coenzymatically activated on posttranslational tethering to a side chain amine of a lysine residue in the approximately 10 kDa carboxyl carrier protein domain of its host carboxylases 18 (Figure 10.7A). The one-carbon molecule carbon dioxide is the minor component in a rapidly established equilibration with HCO₃⁻ and it is bicarbonate that is the proximal substrate for the biotinyl-carboxylases. 2 As anticipated, to activate one of the oxygens of that carbonic acid anion, ATP cleavage via phosphoryl transfer to the oxyanion generates the mixed acyl phosphoric anhydride of carbonic acid and phosphoric acid. This carboxy phosphate (also see Chapter 7) is the activated one-carbon unit that is captured by biotin, perhaps via a transient isoureido species to create the N₁-carboxy-biotinyl-carrier protein domain adduct (Figure 10.7A).

The tethered N₁-carboxybiotin is the coenzyme form that carries the carboxy unit from the ATP-carboxy-phosphate site to a distinct subsite where either the C₂ acetyl carbanion (thioester enolate) of acetyl CoA or the enolate ion form of pyruvate cosubstrates are generated as carbon nucleophiles. These are the kinetically competent nucleophiles for C–COO⁻ bond formation, to give malonyl CoA or oxaloacetate, respectively (Figure 10.7A). All told, there are only half a dozen or so biotin-requiring carboxylases but they are at key metabolic nodal points, reflecting the utility of these CO₂ fixation mechanisms. After almost a century of vitamin B₇ research, the choice of an N-carboxy-urea heterocycle for such metabolic centrality in human metabolism is still surprising.

10.6.2 Folate

Vitamin B₉ was named folate for its high content in folia, the foliage of leafy plants¹. These days, synthetic folate is provided with the bicyclic pterin ring moiety in the air stable fully oxidized form (Figure 10.7B). However, to function coenzymatically, the right-hand pyrazine ring must be reduced twice by two electrons *in vivo*, such that

(5,6,7,8)-tetrahydrofolate is in the tetrahydropyrazine oxidation state. **2** (The left-hand pyrimidine ring stays unchanged.) Also, the monoglutamate tail must typically be elongated to triglutamate to heptaglutamate forms to be recognized by partner enzymes.

The chemical rationale for the tetrahydro oxidation state is that N₅ in the tetrahydropyrazine ring needs to be a basic, nucleophilic nitrogen to function in one-carbon unit transits. Dihydrofolate, formed later in folate metabolic transformations, is the 7,8-dihydro regioisomer, so only in the tetrahydro form is N₅ a basic, secondary amine. The second key basic nitrogen is the amine of the *para*-aminobenzoyl moiety, tethered via a methylene bridge to C₆ of the tetrahydropyrazine ring. Together these two nitrogens (in blue in [Figure 10.7B](#)) will sop up nascent formaldehyde into a latent, strain-free five membered ring system.

The metabolic entry point for billions of one-carbon units in every human cell cycle is the PLP-dependent enzyme serine transhydroxymethylase. This enzyme catalyzes cleavage of the C₃–C₂ bond of L-serine ion the way to glycine and formaldehyde via the L-serine = PLP aldimine intermediate ([Figure 10.7B](#)). Formaldehyde is toxic, more toxic than free ammonia, so cells sponge it up as a latent unreactive derivative attached to tetrahydrofolate. The toxicity of formaldehyde is its indiscriminate reactivity with amine groups, first to form the carbinolamine and then dehydration to the imine. This is even more reactive and captures a second amine. Now, the CH₂ is bridged by two amines. It has crosslinked the two amines. This is the historical use of formaldehyde as an embalming agent. This reactivity of formaldehyde is captured by the N₅ and N₁₀ amino nitrogens in THF that thereby sops up nascent formaldehyde before it can engage in adventitious protein cross-links ([Figure 10.8B](#)).

In the active site of serine transhydroxymethylase the nascent formaldehyde is captured by N₅ of bound THF and dehydrated to the THF–N₅=CH₂ iminium ion. This, in turn, is captured by the adjacent, basic N₁₀-nitrogen to yield the five-membered ring in methylene-THF ([Figure 10.8B](#)). The formaldehyde has crosslinked N⁵ and N¹⁰ and is now situated in a strain-free, unreactive bis-nitrogen heterocycle. This is an efficient and ingenious invention of a physiological formaldehyde sponge or trap. The N⁵,N¹⁰-methylene-THF is then a stable diffusible reagent in cell milieus.

That adduct has three fates. One is to bind to the active site of thymidylate synthase and undergo a net transfer to C₅ of dUMP, in a complex molecular handover that avoids formation of reactive, free formaldehyde. To replicate a three billion base pair genome requires approximately 750 million moles of dTTP per cell cycle, all derivable from the tethered formaldehyde C₁ unit in methylene-THF.

The other two fates are internal redox reactions on N⁵,N¹⁰-CH₂-THF. These are classic NADH–NAD⁺ type amine reductions or oxidations of the tethered formaldehyde equivalent. In the reduction direction an enzyme catalyzes hydride transfer from NAD(P)H to the CH₂ group, perhaps after it has been isomerized back to the CH₂=N₅ iminium ion to give the N₅–CH₃-THF product. This is the immediate donor of the methyl group in biosynthesis of methionine ([Figure 10.8B](#)), one of the 20 amino acid building blocks required for protein biosynthesis. In the opposite, oxidative direction the N⁵,N¹⁰-CH₂-THF gives up one of the two hydrogens as a hydride group to cosubstrate NAD⁺. The products are NADH and the N⁵,N¹⁰-methenyl cyclic adduct. That can be opened by a hydrolytic enzyme to N¹⁰-formyl-THF. It is the immediate donor of formyl groups at two stages of purine ring biogenesis ([Figure 10.8B](#); purine biosynthesis logic is examined in Chapter 9).

Overall, THF with its pair of basic amines at N⁵ and N¹⁰ is an ingeniously evolved organic scaffold to act as a two-pronged nitrogen sponge for nascent formaldehyde to avoid letting it go free to react with cellular constituents. Yet, that formaldehyde equivalent is crucial for DNA, purine and methionine biosynthesis.

THF can thus act as a specific one-carbon donor at the original oxidation state of

formaldehyde (thymidylate synthase), at the two-electron reduced state of methanol (methionine synthase), or at the two-electron oxidized state of formate (purine ring assembly).

As in the one-carbon metabolite cyanide ion, analyzed in Chapter 13, the shuttling of tethered one carbon units, CO_2 , HC=O , $\text{H}_2\text{C=O}$, and $\text{H}_3\text{C-OH}$ equivalents between nitrogenous cosubstrates, reveal fundamental nitrogen chemical biology. Between them, the nitrogen centers in biotin and tetrahydrofolate frameworks are able to ferry the whole range of one-carbon oxidation states between organic scaffolds in essential biosynthetic processes.

10.7 Vitamin B₁₂: Two Organometallic Coenzyme Forms in a Modified Tetrapyrrole Macrocycle

The last of the B vitamins that function as coenzymes is vitamin B₁₂, one of the metallo-tetrapyrrole macrocycles discussed in Chapter 11. 1.2 Its biosynthesis from the typical tetrapyrrole macrocyclic framework is discussed in that chapter.

The top axial ligand to the central cobalt^{III} ion can be any of four substituents. 19 Vitamin B₁₂ as isolated from natural sources can be in the aquo form with water as the top ligand (Figure 10.9A). Or, it can have picked up an adventitious equivalent of a cyanide ion to give cyanocobalamin, an artifact of the isolation procedure (see Chapter 13 for nitrile and cyanide ion chemical biology).

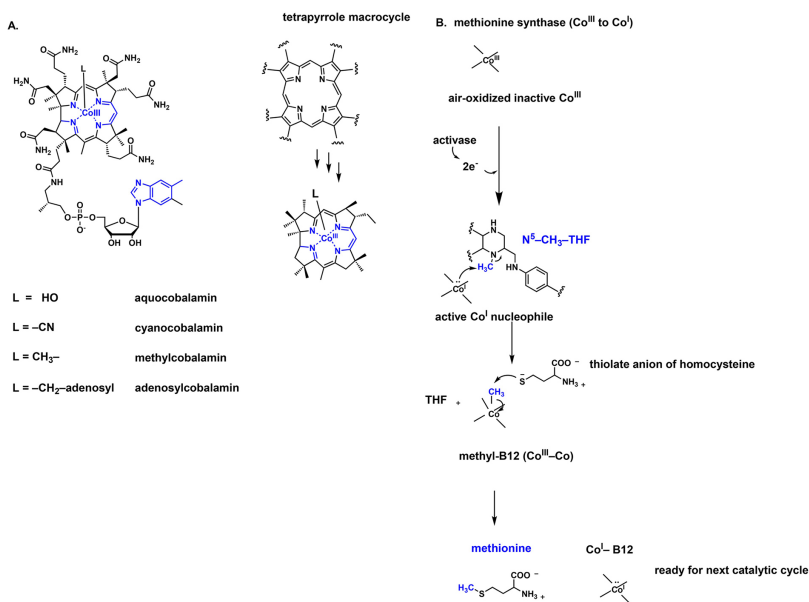


Figure 10.9 Vitamin B₁₂ and its coenzyme forms. (A) A ring-contracted tetrapyrrole macrocycle with a cobalt ion in the equatorial plane of the tetrapyrrole ring. The pendant dimethylbenzimidazole is the bottom axial ligand to cobalt. (B) The cobalt ion in B₁₂ cycles between Co^{III} and Co^{I} in methionine synthase.

When the top axial ligands are carbon atoms, two organocobalt linkages, unusual organometallic species in biology, result. Methyl B₁₂ is a transient intermediate in the human B₁₂-dependent methionine synthase (Figure 10.9B). By comparison, adenosyl-B₁₂, with its cobalt–C5' bond to the ribose ring of adenosine is a stable entity and the resting form in the human enzyme methylmalonyl CoA mutase.

In methionine synthase catalysis the equatorial cobalt ion in the corrin ring cycles between Co¹⁺ and Co³⁺ and back to Co¹⁺ in each catalytic cycle. In contrast the adenosyl B₁₂ coenzyme undergoes homolysis of the Co–C5' bond to generate Co²⁺ and a 5'-deoxyadenosyl radical [dA•] in each catalytic cycle (Figure 10.10A). The [dA•] serves as radical initiator to abstract a hydrogen atom [H•] in a bound cosubstrate. The resultant substrate radicals rearrange to product radicals, then obtain an [H•] back from dA–H. The final step is recombination of the [dA•] radical and the Co^{II} form of B₁₂ to regenerate adenosyl B₁₂ in time for a subsequent catalytic cycle.

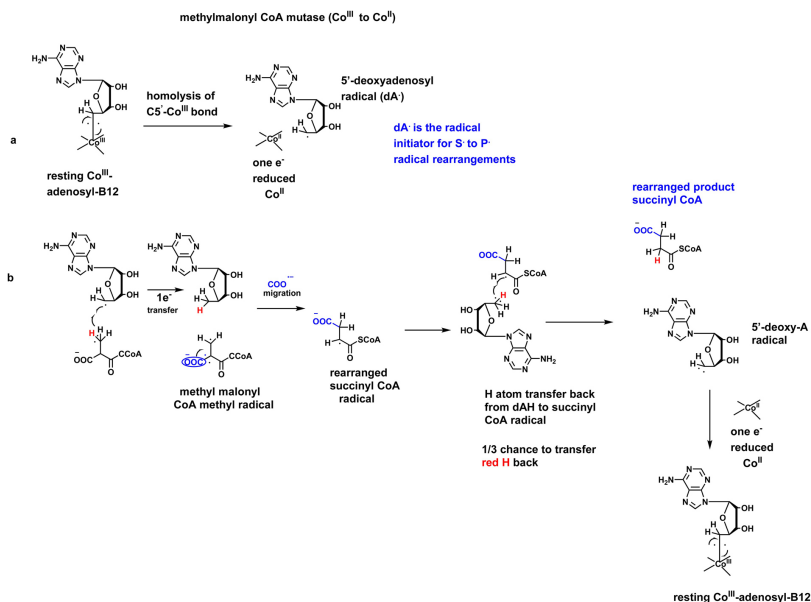


Figure 10.10 The cobalt ion in B₁₂ cycles between Co^{III} and Co^{II} with radical intermediates in the methylmalonyl CoA mutase reaction. Step a shows the homolytic fragmentation of the Co^{III}–CH₂ bond of adenosyl B₁₂ to Co^{II} and the 5'-deoxyadenosyl radical [dA•]. In step b that deoxyadenosyl radical is the initiator of substrate to product radical rearrangements by abstraction of an [H•] from bound substrate.

The first logic of two-electron cycling of the B₁₂ cobalt ion between Co^I and Co^{III} is exemplified in human methionine synthase. The enzyme as isolated aerobically must be incubated with a reducing system to convert the air-oxidized Co³⁺ state back to the Co¹⁺ state that is the active nucleophilic form of cobalt. The methyl donor is the N⁵-methyl THF (see Figure 10.7) and it is captured by the supernucleophilic Co^I to release THF and form the organometallic methyl-B₁₂ intermediate (Figure 10.9B). Note that the [CH₃] group is transferred as a [CH₃]⁺ species. The cobalt in the resultant methyl B₁₂ thus has the formal oxidation state of Co³⁺.

The second substrate in the methionine synthase reaction is the thiolate anion of L-homocysteine. The thiolate nucleophile attacks the CH₃ group in the top axial position

of methyl-B₁₂. Again, the itinerant CH₃ group is moving to sulfur as a [CH₃]⁺ equivalent. Therefore, the departing cobalt ion leaves with the electron pair that had been in the Co–CH₃ bond and so returns to its nucleophilic Co^I oxidation state at the end of the catalytic cycle. This is a cobalt ion–organocobalt two-electron reaction manifold.

In contrast, the adenosyl B₁₂ coenzyme in the human enzyme methylmalonyl CoA mutase reacts in a one-electron manifold, cycling between Co^{III} in the starting adenosyl B₁₂ and Co^{II} during the radical transfers in each catalytic cycle. The adenosyl-B₁₂ serves as a radical generator, undergoing homolysis of the Co–C bond in the enzyme active site to give the 5'-deoxyadenosyl radical (dA) that is a robust oxidant (Figure 10.10A) and a radical initiator of cosubstrate radicals. After the initial homolysis the Co^{II} form of B₁₂ is just a bystander till the end of the reaction.

The incipient dA' radical abstracts a hydrogen atom from nearby C–H bonds, even if those bonds are unactivated. In the case shown (Figure 10.10A), one of the three methyl hydrogens of bound methylmalonyl CoA is abstracted as [H]. This generates dA–H (deactivated) and the methyl radical on the methylmalonyl scaffold. The kinetically competent fate is homolysis of the C–COO bond in that molecule with migration of the [COO[–]] group to the initial radical site. This constitutes a radical-based skeletal rearrangement. The initial branched chain carbon framework is now a straight chain framework, with radical density at the indicated [CH] carbon (succinyl CoA carbon radical). This product-like radical can abstract an H atom [H] back from the dA–H. This produces the succinyl CoA product and regenerates the [dA] radical. Now, the bystanding Co^{II}–B₁₂ engages in radical recombination with [dA] to reform the starting adenosyl B₁₂ (cobalt oxidation state of Co^{III}).

The two human enzymes, methionine synthase and methylmalonyl CoA mutase display the two dramatically different reaction manifolds that B₁₂ coenzymes can undergo. Together, the two-electron and one-electron pathways show the full range of the three oxidation states of the cobalt ion (Co^{I–III}) in the corrin modified tetrapyrrolic macrocycle. The corrin macrocycle is an encasement of a cobalt ion just barely manageable in terms of adventitious reactivity in the cellular milieu, segregated to anaerobic microenvironments of aerobic cells.

10.8 Overview

The scope of enzyme-mediated chemical transformations effected by the nitrogen heterocycles of the seven B vitamins radiates through the many branches of primary metabolism. The coenzyme forms of each of these B vitamins uses a different nitrogen functional group, illustrating the diaspora of nitrogen heterocyclic chemistry and indicating that they were early essential players in molecular evolution. Along their own evolutionary path, humans lost the ability to biosynthesize these coenzyme scaffolds but have not found any replacements for them. Hence, their essential vitamin status.

The diversity of these vitamin-based nitrogen heterocycles and their artful biosynthetic construction leads to accessible chemistry not available to the 20 amino acid side chains in proteins. The vitamins as coenzymes vastly expand nature's catalytic inventory.

For example, the biosynthetic quaternization of the thiazole ring in vitamin B₁ and of the pyridine ring in maturation of niacin to NAD(P)⁺ not only lead to permanently charged nitrogens but also enable the specific and distinct chemistries: the C₂ thiazolium ylide in thiamin pyrophosphate *versus* the hydride transfer to C₄ of the pyridinium ring of NADP⁺. In contrast, the pyridine ring in pyridoxal phosphate is uncharged and as a neutral heterocycle conjugated to the external aldehyde group, the

pyridine ring can act as an electron sink to stabilize amino acid- substrate carbanions. Thus, vitamins B₁ and B₆ use nitrogen heterocycles to stabilize carbanionic transition states while, conversely, NAD(P)⁺ (from vitamin B₃) mediates hydride ion transfers to and from dozens of carbonyl group substrates.

The bicyclic 5,5-fused scaffold of biotin (B₇) and the N⁵,N¹⁰-framework of tetrahydrofolate (B₉) serve to carry one-carbon metabolic fragments at four different oxidation states as latently reactive one-carbon adducts, deliverable in the active sites of target enzymes for biosynthetic carboxylations, aldol condensations, methylations and formylations in diverse but central biosynthetic pathways.

The tricyclic isoalloxazine ring system of flavin coenzymes (vitamin B₂) puts them in a class uniquely versatile as redox reagents in biology. N₅ can be electrophilic in the oxidized state, nucleophilic in the dihydroflavin reduced state and a hydrogen atom donor with resultant N₅ radical in the semiquinone radical state. The ability of the nitrogen heterocyclic framework to serve as both two-electron and one-electron redox switch allows flavins to function at multiple such stepdown electron transport intersections. This property also serves as a transition to Chapter 14 on one-electron chemical biology.

References

1. G. F. Combs and J. McClung, *The Vitamins: Fundamental Aspects in Nutrition and Health*, Academic Press, 5th edn, 2017.
2. C. T. Walsh and Y. Tang, *The Chemical Biology of Human Vitamins*, Royal Society of Chemistry, London, 2019, p. 446.
3. C. T. Walsh Nature loves nitrogen heterocycles, *Tetrahedron Lett.*, 2015, **56**, 3075-3081.
4. C. Badii and L. Reed-Guy, BeriBeri: from <https://www.healthline.com/health/beriberi>, 2017.
5. R. Breslow Rapid deuterium exchange in thiazolium salts, *J. Am. Chem. Soc.*, 1957, **79**, 1762-1763.
6. K. Tittman, K. Mesch, M. Pohl and G. Hubner Activation of thiamin diphosphate and FAD in the phosphate dependent pyruvate oxidase from *Lactobacillus plantarum*, *J. Biol. Chem.*, 1998, **273**, 12929-12934.
7. C. T. Walsh and T. Wencenwicz Flavoenzymes: versatile catalysts in biosynthetic pathways, *Nat. Prod. Rep.*, 2013, **30**, 175-200.
8. S. Adak and T. Begley Flavin-N⁵-oxide: a new motif in flavoenzymology, *Arch. Biochem. Biophys.*, 2017, **632**, 4-10.
9. A. Nikiforov, V. Kulikova and M. Zeigler The human NAD metabolome: functions, metabolism and compartmentalization, *Crit. Rev. Biochem. Mol. Biol.*, 2015, **50**, 284-297.
10. C. T. Walsh, B. P. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
11. H. F. Fisher, E. E. Conn, B. Vennesland and F. H. Westheimer The enzymatic transfer of hydrogen: the reaction catalyzed by alcohol dehydrogenase, *J. Biol. Chem.*, 1953, **202**, 687-697.
12. Y. Kwan-sa, L. J. Arnold, W. S. Allison and N. O. Kaplan Enzyme stereospecificities for nicotinamide coenzyme, *Trends Biochem. Sci.*, 1978, **3**, 265-268.
13. I. J. Ozaki and J. Moss Structure and function of the mono-ADP-ribosyltransferases, *Rev. Physiol., Biochem. Pharmacol.*, 1996, **129**, 51-94.
14. M. Mashimo, J. Kato and J. Moss Structure and function of the ARH family of ADP ribose acceptor hydrolases, *DNA Repair*, 2014, **2014**(0), 88-94.
15. P. K. Mehta and P. Christen The molecular evolution of pyridoxal-phosphate-

- dependent enzymes, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1999, **74**, 129-184.
16. H. C. Dunathan Conformation and reaction specificity in pyridoxal phosphate enzymes, *Proc. Natl. Acad. Sci. U. S. A.*, 1966, **55**, 712-716.
17. R. A. John Pyridoxal-P dependent enzymes, *Biochim. Biophys. Acta*, 1995, **1248**, 81-96.
18. L. Tong Structure and function of biotin-dependent carboxylases, *Cell. Mol. Life Sci.*, 2013, **70**, 863-891.
19. R. Banerjee and S. W. Ragsdale The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes, *Annu. Rev. Biochem.*, 2003, **72**, 209-247.

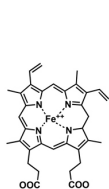
Tetrapyrrole Macrocycles: Aminopyrroles and the Pigments of Life

11.1 Four Classes of Metallated Tetrapyrrole Macrocycles

In Chapters 2–5 on inorganic nitrogen metabolism, particularly in nitrification, denitrification and anammox arms of the biogeochemical nitrogen cycle, the redox active enzymes repeatedly used heme-type tetrapyrrole macrocycles as electron conduits and active sites of electron input (reduction) or exit [oxidation of the inorganic nitrogen substrates and intermediates (Figure 5.5)]. Iron-containing tetrapyrrole macrocycles were early inventions enabling lithotrophic microorganisms to eke out a living oxidizing and reducing inorganic nitrogen compounds.

These pyrrolic macrocycles have persisted indispensably throughout evolution, with radiation to four distinct types of tetrapyrrole frameworks, based on the metal cation held in the equatorial plane of the four rings **1** (Figure 11.1). Sometimes termed the pigments of life, the iron atom (Fe^{II} or Fe^{III}) in the tetrapyrrole equatorial plane of heme imparts a red–brown color to the molecule with classic absorption maximum at 410–515 nM. Also shown are the structures of siroheme, the violet-colored cofactor in nitrite reductase (Chapter 4) and the greenish linear tetrapyrrole termed bilane (discussed in Section 11.2.3).

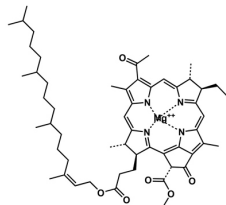
heme b (ferrous iron)



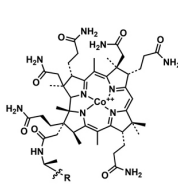
red-brown

(dioxxygen carrier or
electron transport)

bacteriochlorophyll a



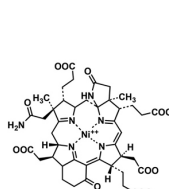
green

Mg²⁺ metallocycle
(photosynthesis)vitamin B₁₂

reddish pink

Co³⁺ metallocycle

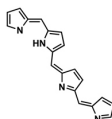
Coenzyme F430



yellow

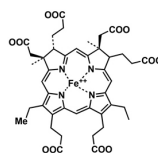
Ni²⁺ metallocycle

Bilane



(linear tetrapyrrole)

Siroheme



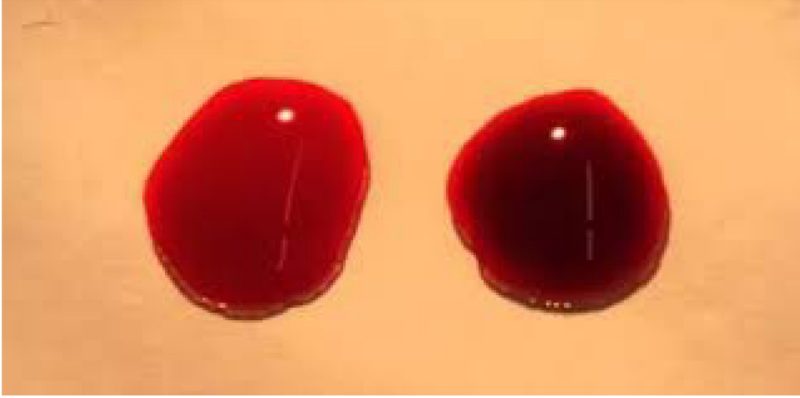
purple-violet

Figure 11.1 Tetrapyrroles: heme B (Fe^{II}) is the initial biosynthetic iron-metallated porphyrin and is red; bacteriochlorophyll with Mg²⁺ as the cation chelated in the equatorial plane of the macrocycle is green and has a long membrane anchor; vitamin B₁₂ the Co^{III} tetrapyrrole macrocycle is reddish-pink; nickel-containing coenzyme F430 in methane biogenesis is yellow; siroheme in nitrite reductase has eight carboxy substituents and is violet colored; the linear tetrapyrrole is termed a bilane.

The heme cofactor in the hemoglobin of red cells is the reason those cells are red (Figure 11.2). Oxy-hemoglobin at high [pO₂] levels has the bright red color of arterial blood, while deoxygenated hemoglobin, after its passage through tissues far from the lungs, has the wine dark red color.

**The red color of drops of blood
is due to the heme in red cell hemoglobin**

**Oxy-Hemoglobin in arterial blood on the left
Deoxy-Hemoglobin from vein pus blood on the right**



**The green color of grass is due to
chlorophyll, reflecting back green light**



Figure 11.2 The pigments of life: blood is red because of hemoglobin with different tones for oxy- and deoxy-hemoglobin. Reproduced from <https://commons.wikimedia.org/w/index.php?curid=8543382>, under the terms of the CC BY 3.0 license, <https://creativecommons.org/licenses/by/3.0/deed.en>, credit: <http://www.biophotonicsworld.org/uploads/43>. All chlorophyll containing plants are green. Reproduced from https://commons.wikimedia.org/wiki/File:M%C3%A9lisse_Feuilles_FR_2013b.jpg, under the terms of the CC BY-SA 3.0 license <https://creativecommons.org/licenses/by-sa/3.0/>, credit: JLPC.

When Mg^{2+} is the central ion in the chlorophyll macrocycles (chlorophyll a or b) there is an absorbance maximum at 660–680 nM. In fact, chlorophylls absorb both red

and blue light, reflecting green light. The green plant life in [Figure 11.2](#) is all due to green chlorophylls. Less widely distributed are the salmon-colored pigment vitamin B12, with equatorial cobalt^{III} ion as isolated and the yellow nickel^{II}-containing coenzyme F430 of anaerobic methanogenic bacteria.

11.2 Heme Biosynthesis

Essentially all free-living organisms have retained the ability to biosynthesize tetrapyrrole macrocycles. [1.2](#) Although free pyrrole is not a known biological metabolite, pyrrole 2-carboxylic acids and the trisubstituted pyrrole known as porphobilinogen are naturally occurring. The pyrrole-2-carboxylates are usually microbial metabolites derived from net four-electron oxidation of the common amino acid L-proline ([Figure 11.3](#)).

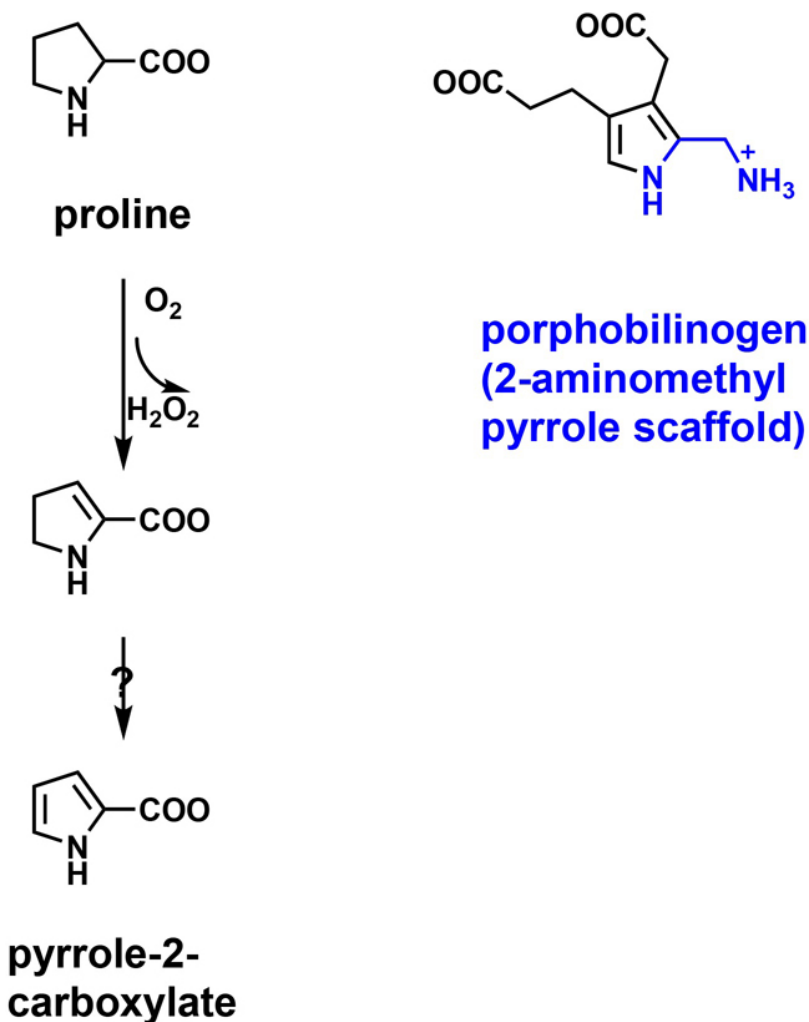


Figure 11.3 Two routes to pyrrole metabolites: proline to pyrrole-2-carboxylate in

microbial metabolism; porphobilinogen, a 2-aminomethylpyrrole, is the central building block for all tetrapyrrole macrocycles.

The trisubstituted porphobilinogen (Figures 1.1 and 11.3) that is the precursor to all tetrapyrrole macrocycles is instead assembled in turn from the unusual 5-amino-4-ketopentanoate, known as aminolevulinate (ALA) by distinct logic among different phyla 2 (Figure 11.4). The two molecules aminolevulinate and porphobilinogen are the key structures in the first part of the biosynthetic route to the tetrapyrrolic pigments of life.

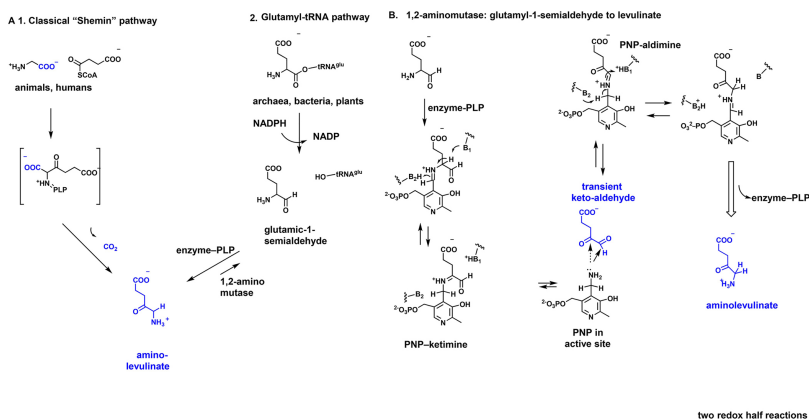


Figure 11.4 (A) Two routes to the 4-keto-5-amino acid aminolevulinate. The Shemin pathway in humans condenses glycine and succinyl CoA via a PLP enzyme that is both a decarboxylase and a C–C bond synthase. Microbes and plants instead reduce glutamyl-tRNA at the C₁-oxoester to free glutamyl-1-semialdehyde. A PLP-dependent 1, 2-aminomutase converges on amino-levulinate; (B), the 1,2-aminomutase catalyzes an unusual internal transfer of the –NH₂ group from C₂ to C₁ by way of the reduced PNP coenzyme and a nascent 1,2-ketoaldehyde. The numbering changes in levulinate from a 2-keto-1-amino- product to numbering from the remaining carboxylate make aminolevulinate a 4-keto-5-amino-valerate.

11.2.1 Two Routes to Aminolevulinate

In archaea, bacteria, and plants ALA is obtained from glutamate when the C₁ carboxyl group is activated as glutamyl-tRNA^{Glu}, the metabolite normally used for protein biosynthesis (Figure 11.4A). 3 The oxoester bond between the glutamate C₁ carbonyl and the 3'-OH of the AMP residue on the tRNA is reduced by an NADPH-dependent reductase, releasing the glutamyl fragment as the C₁-aldehyde, known as glutamyl-1-semialdehyde. The next enzyme contains one equivalent of pyridoxal phosphate and carries out a remarkable 1,2-isomerization, in which the C₁ carbonyl and the C₂-amine groups are switched, yielding ALA (the numbering is now from the C₁-carboxylate) (Figure 11.4B). On examination of mechanism, this isomerase is an intramolecular variant of a transaminase, forming the two-electron reduced pyridoxamine phosphate (PNP) and the dicarbonyl intermediate (a keto-aldehyde). Before that nascent keto-aldehyde dissociates, the enzyme catalyzes the second half reaction, reoxidation of the PNP and reductive amination of the dicarbonyl framework of the intermediate. Reduction of the ketone would go back to the starting glutamate semialdehyde. Reduction of the aldehyde is the forward direction, creating the 4-keto-5-amine constellation in released aminolevulinate.

Animals, including humans, have discarded this approach and instead start with the

common metabolites glycine and succinyl CoA. This pathway was elucidated largely by the efforts of David Shemin and is often referred to as the Shemin pathway 2 (Figure 11.4A). The aminolevulinate synthase is another example of amino acid-dependent pyridoxal phosphate scaffold manipulation. Comparing the two PLP-enzymes in the two pathways with ALA (Figure 11.4A) gives insight into the many ways PLP, directed by the conformations imposed by their respective apoproteins, utilize aldimine-enabled chemistry to modify amino acid frameworks.

In ALA synthase a stabilized glycyl-PLP C₂ carbanion attacks the electrophilic thioester carbonyl of succinyl CoA to generate a branched six-carbon diacid adduct (Figure 11.5). Then, decarboxylation of what had been the glycine carboxylate is a low energy pathway to generate a second stabilized PLP carbanion. Protonation at what is now numbered as C₅ of the product leads to release of aminolevulinate. Of the starting two-carbon amino acid glycine, the C₁ carboxylate has been discarded as CO₂ and the C₂ carbon (CH₂NH₃) has become appended to what had been the electrophilic thioester carbonyl of succinyl CoA. The structure of ALA, generated by either of the two routes in Figure 11.4A, is unusual in primary metabolism, with the amine group part of a keto-amine moiety tacked on to the succinate scaffold.

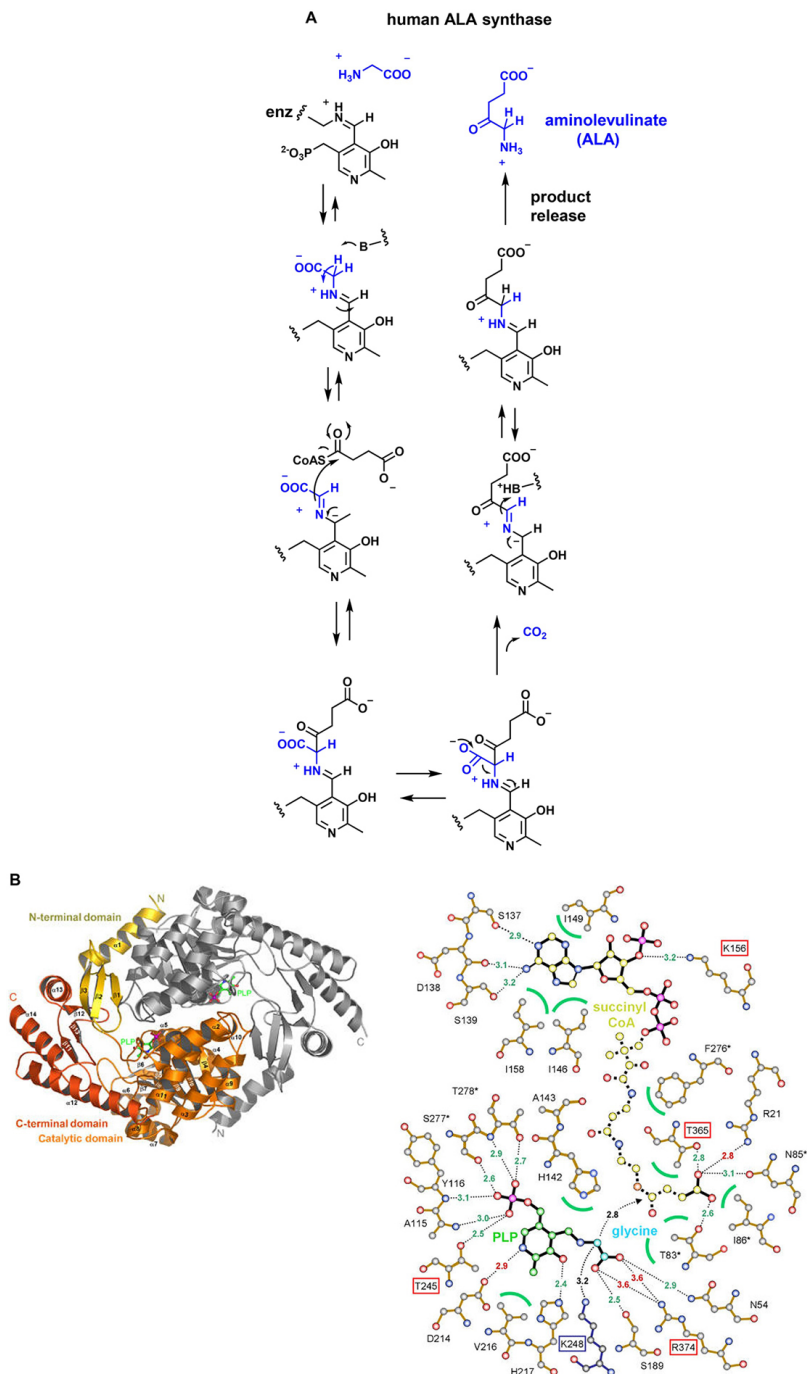


Figure 11.5 (A) Mechanism of the condensation of glycine with succinyl CoA by the PLP–enzyme aminolevulinate synthase. (B) Architecture of the Ala synthase dimers and interactions of enzyme side chains with both succinyl-CoA and glycyl-PLP. Orientation of ALA synthase dimer with catalytic domain in C-terminus of

each. Reproduced from [ref. 22](#) with permission from John Wiley & Sons, Copyright © 2005 European Molecular Biology Organisation. Adapted from [ref. 23](#) with permission from Elsevier, Copyright 1998. Interactions of both succinyl-CoA and the glycyl-PLP adduct in the ALA synthase active site. Reproduced from [ref. 22](#) with permission from John Wiley & Sons, Copyright © 2005 European Molecular Biology Organisation.

11.2.2 Aminolevulinate to Porphobilinogen: Imine and Enamine Chemistry

The next enzyme in the pathway is porphobilinogen synthase. [4](#) It is the key catalyst creating the aminomethylpyrrole framework that is the essential structural element for oligomerization to tetrapyrroles. The adjacency of the C₅ amine and C₄ ketone is the key to enzymatic intermolecular dimerization and double dehydration of two molecules of aminolevulinate to produce the trisubstituted heteroaromatic pyrrole monomer. A pair of linear amino ketone moieties have been morphed into the aromatic five-membered pyrrole heterocycle.

When the dust has settled, one C–C and one C–N bond and two dehydration events have occurred, yielding a trisubstituted pyrrole bearing a CH₂–NH₂ group at C₂, an acetic acid chain at C₃ and a propionic acid chain at C₄. This molecule is known as porphobilinogen, literally a generator of the porphyrin and bilane scaffold (see [Figure 11.1](#) for bilane structure). The different length acetic and propionic acid chains at adjacent carbons of the pyrrole ring arise from the “out of register” C–C bond forming steps between the two 5-ALA monomers. The enzyme uses both imine and enamine chemistry, aided by a catalytic lysine side chain. Imine chemistry creates a more electrophilic carbonyl group in one of the ALA molecules. Tautomerization of iminium to enamine creates the carbanion equivalent needed for C–C bond formation ([Figure 11.6](#)).

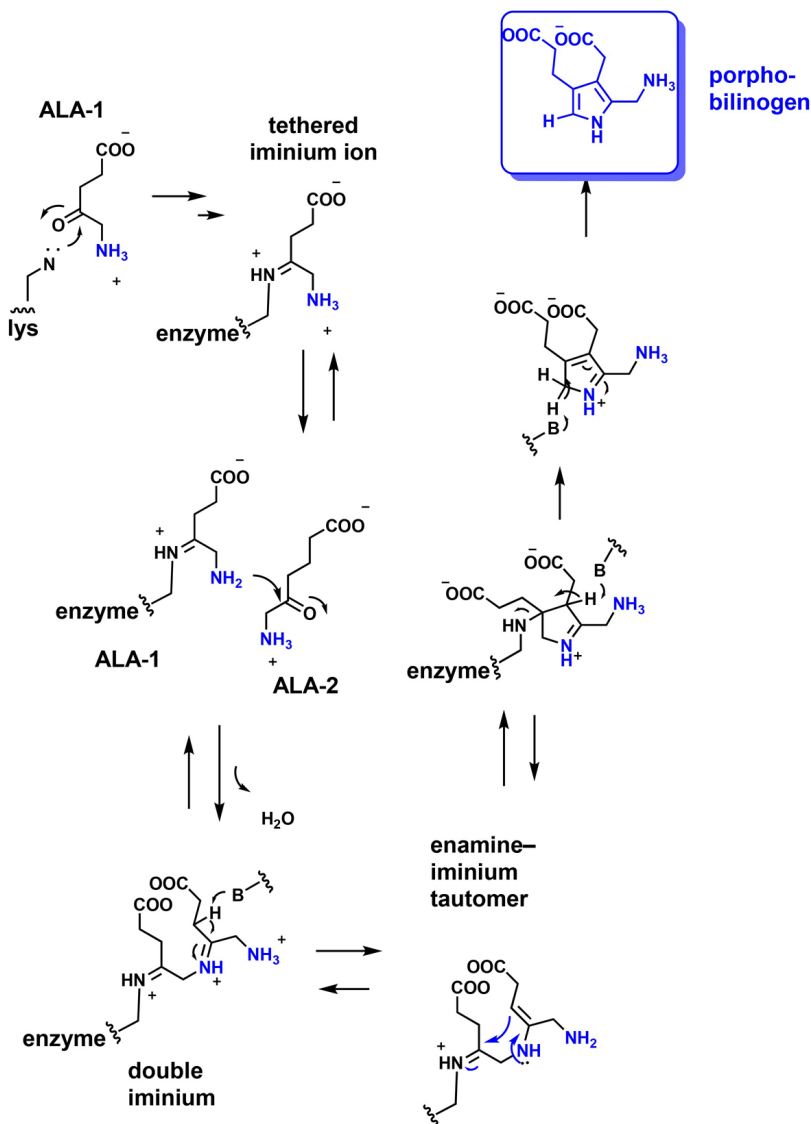


Figure 11.6 Porphobilinogen synthase: dimerization of a pair of aminolevulinates via iminium ion–enamine chemistry: two C–C bonds made; two water molecules lost.

11.2.3 Porphobilinogen

The 2-aminopyrrole scaffold of porphobilinogen is at the heart of the pathway logic to all four classes of metallated tetrapyrroles **1** (Figure 11.1). The pyrrole ring is aromatic with delocalization of the nitrogen lone pair into the π -electron cloud of the two C–C double bonds. Thus, the pyrrole nitrogen is not notably basic or nucleophilic. However, it can enable accessible, low energy C₂ carbanion formation (Figure 11.7). That propensity in the 2-aminomethyl pyrrole framework of porphobilinogen leads to facile expulsion of free ammonia with formation of an exocyclic double bond (lines 2 and 3

of Figure 11.7). This exocyclic methylene is an *azafulvene* intermediate that underlies enzymatic construction of a linear tetrapyrrolic intermediate. Azafulvenes are electrophilic at the terminus of the exocyclic double bond (shown in the resonance contributor), as electron flow from a nucleophilic attack restores the pyrrole double bond configuration.

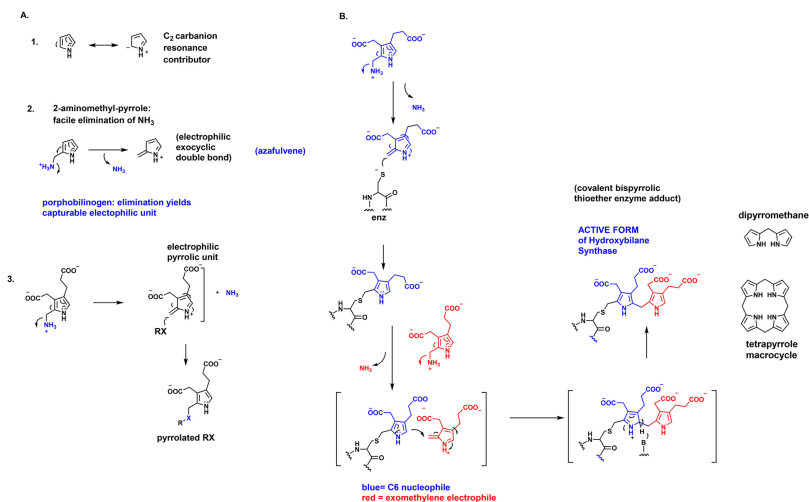


Figure 11.7 Pyrrole ring reactivity: (A) C₂-carbanion character in pyrrole ring; electrophilic exocyclic double bond (azafulvene) structure from facile elimination of side chain NH₃ from 2-aminomethyl pyrroles; capture of that exocyclic methylene by a nucleophile RX. (B) Mechanism for autoactivation of hydroxybilane synthase as covalent bi-pyrrolic thioether enzyme adduct by a combination of C₆ as nucleophile and azafulvene structure as electrophile in pairs of porphobilinogen monomers; nomenclature and structure for dipyrromethane vs. tetrapyrrole macrocycle.

In the next section on hydroxybilane synthase we will observe that the aminomethylpyrrole moiety in porphobilinogen in elimination of NH₃ to form the azafulvene structure creates that olefinic electrophile. A second molecule of the pyrrole can capture the azafulvene by acting as a C₆ carbon nucleophile. The ability of the aminomethylpyrrole intermediates to act *iteratively as electrophiles then nucleophiles* is a reaction duality put to use to construct the tetrapyrrole macrocycles (Figure 11.7B).

11.2.4 Porphobilinogen To Hydroxybilane

The third enzyme in the human pathway to heme is hydroxybilane synthase, **5** converting the monocyclic bilane precursor porphobilinogen to the linear tetrapyrrolic hydroxybilane, an unstable molecule with an estimated *in vivo* half time of approximately 4 minutes. This enzyme has also been termed porphobilinogen deaminase, to emphasize that it is the step that releases four equivalents of NH₃ in the tetrapyrrole biosynthetic pathway.

The dual electrophilicity–nucleophilicity of the monopyrrolic aminomethyl framework is utilized multiple times by this enzyme. The pyrrole moiety in porphobilinogen, with its C₂–CH₂NH₂ substituent, can act either as an electrophile at the exocyclic CH₂ group of the azafulvene or as a nucleophile at C₆, set up for iterative C–C bond forming chain extensions.

However, first the enzyme primes itself from inactive to active form by covalently tethering two pyrroles as a methylene bridged covalent dipyrrole in the enzyme's active site. ¹ This is a unique *auto-posttranslational* modification process, thought to start with a cysteine thiolate anion in the apoenzyme capturing the azafulvene electrophilic exomethylene (Figure 11.7A) derived from porphobilinogen (blue color in Figure 11.7B). ⁶ This is a covalent monopyrrolic adduct (a pyrromethane), a way station between inactive apoenzyme and active dipyrrolic enzyme. (The dipyrrole moiety has been termed a *dipyrromethane* group: as shown in Figures 11.7B and 11.8, it resembles one half of a tetrapyrrole macrocycle.)

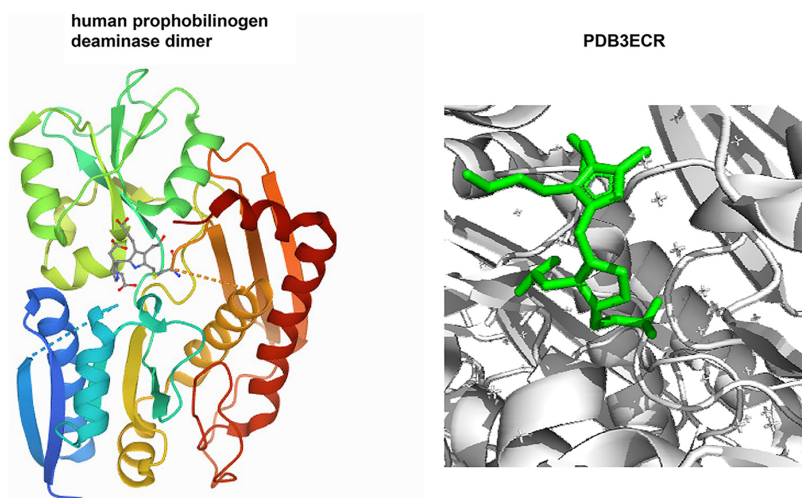


Figure 11.8 Ribbon diagram for human porphobilinogen deaminase, with close up view of the dipyrromethane adduct in the primed enzyme, from PDB file 3DECR.

That covalently tethered monopyrrole ring acts as C₆ nucleophile towards the azomethine form of a second porphobilinogen (shown in red). A new C–C bond forms between C₆ of the blue pyrrole to the exomethylene carbon of the red azomethine tautomer of the pyrrole. After rearomatization of the blue pyrrole, the hydroxybilane synthase has arrived at its unique active state. ⁶

Startlingly, the active hydroxybilane synthase then proceeds to iterate the C–C bond formation four more times with porphobilinogen substrate monomers that have eliminated NH₃ to generate the electrophilic azomethine reactant partners. This part of the catalytic cycle ends with accumulation of a hexapyrrolic adduct in the active site (Figure 11.9). In each enzyme-mediated elongation event the incoming porphobilinogen monomer acts as the electrophile after loss of the side chain amine, while the unit added in the previous step acts as the nucleophile.

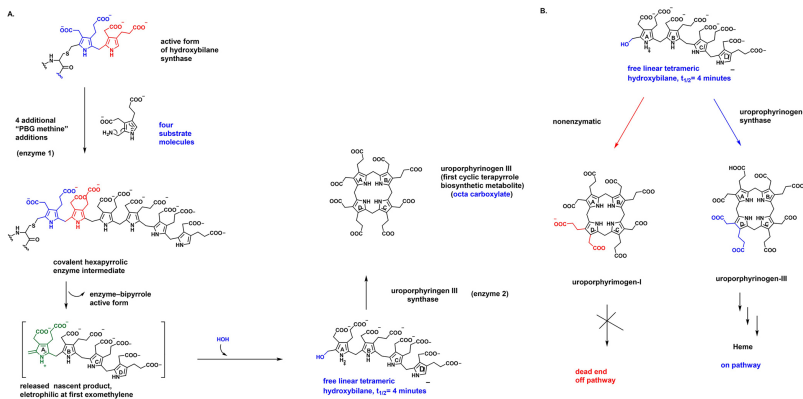


Figure 11.9 (A) Four porphobilinogen monomers give rise to first cyclic “on pathway” intermediate uroporphyrinogen III. Hydroxybilane synthase acts *via* a covalent hexapyrrolic enzyme intermediate that undergoes regiospecific hydrolysis to liberate the linear tetrapyrrolic hydroxybilane. (B) Hydroxybilane is labile to nonenzymatic cyclization to give the symmetric uroporphyrinogen I. The enzyme uroporphyrinogen III synthase instead directs the cyclization to the unsymmetric macrocycle uroporphyrinogen III.

Then, perhaps controlled by volume of the accumulating bilane chain, the enzyme catalyzes fragmentation between the second and third pyrrolic residues, returning the enzyme to its active dipyrrolic covalent adduct, ready for the next catalytic cycle (Figure 11.9). The product released is the nascent linear tetrapyrrole chain (bilane) with an electrophilic azomethine moiety in the indicated A ring. That electrophilic exomethylene group is captured by a water molecule to yield the accumulating released linear tetrapyrrole hydroxybilane, accumulating for only a few minutes before it would react off pathway.

The nascent, inactive hydroxybilane synthase from scratch builds itself into a covalent dipyrrolic adduct (dipyrromethane scaffold) that then goes on to add four more pyrrole monomers by the same nucleophile–electrophile pairing to create six methylene bridged pyrrole units. At the maximum length of six tethered pyrroles, the enzyme regiospecifically induces a retro-fragmentation between pyrroles 2 and 3. All in all, a remarkable elongation strategy to build the tetrapyrrolic bilane scaffold, enabled by the pyrrole heterocyclic nitrogen chemistry.

Inter alia, the *head to tail* elongation chemistry for tetrapyrrole assembly by the deaminase–hydroxybilane synthase, is reminiscent of prenyl transferases that use one isomer of isopentenyl-PP as an electrophile (Δ^2) and the other double bond isomer (Δ^3) as a carbon nucleophile in head to tail chain elongations. In this case it is the versatility of the aminopyrrole porphobilinogen building blocks to act alternately as a nucleophile and an electrophile that reveals the chemical versatility of this five-member nitrogen heterocycle for these biological assembly roles.

11.2.5 Uroporphyrinogen-III Synthase

If left to its own devices, the hydroxybilane would rapidly cyclize to a symmetrically substituted macrocycle known as uroporphyrinogen-I (Figure 11.9B). That tetrapyrrole macrocycle is off pathway and not useful in any of the subsequent metallated tetrapyrrole biosyntheses. The purpose of uroporphyrinogen-III synthase is to capture the unstable hydroxybilane before its nonenzymatic cyclization and instead vector it to the isomer known as uroporphyrinogen-III 7 (Figure 11.9B).

It has been observed that humans with defective forms of uroporphyrinogen III synthase, accumulate that distinct, inactive tetracyclic product uroporphyrinogen I (Figure 11.9B). This is the agent that on buildup in various tissues causes the pathological symptoms of acute intermittent porphyria. 8

This uroporphyrinogen-III synthase carries out another remarkable chemical transformation in a pathway full of them, with the pyrrole rings showing off their chemical virtuosity. The synthase flips the orientation of the D ring with respect to the A and C rings as ring closure is effected. 1 As suggested in Figure 11.10, the synthase accelerates cyclization of a folded conformer of linear substrate hydroxybilane. Loss of water gives the azafulvene form of pyrrole ring A that is captured as electrophile by C₂ of ring D acting as nucleophile. That regiochemical control of electron flow would give a spiro adduct (spiropyrrolenine) between rings A and D. That can undergo fragmentation with reverse polarity to an azomethine in ring D and a nucleophilic pyrrole in ring A. Reformation of the methylene bridge between rings A and D leads to the uroporphyrinogen III regiochemistry with asymmetrically disposed acetate and propionate side chains around the macrocycle, imposed by the flip in orientation between the D ring and A ring. No direct evidence for a spiro intermediate has been accumulated, but it offers a reasonable path to the observed uroporphyrinogen-III regiochemistry.

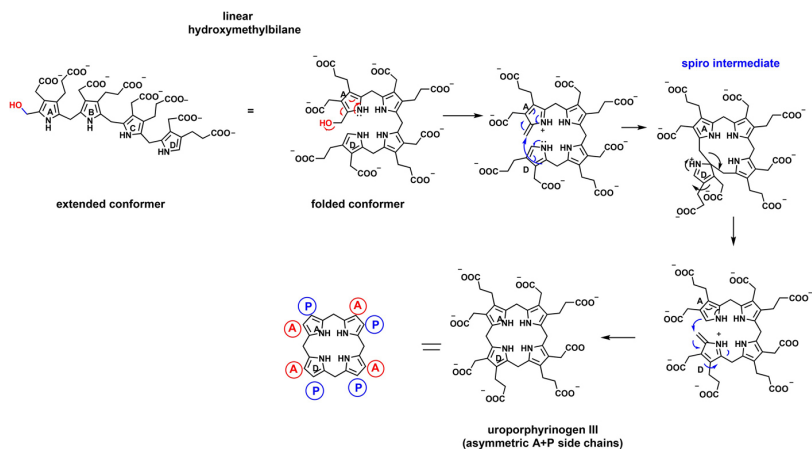


Figure 11.10 In uroporphyrinogen III the D ring has been flipped during the enzymatic process of cyclization to change the orientation of the acetate and propionate side chains. A proposed spiro intermediate utilizes the chemistry of pyrroles in a pair of additions and eliminations.

As a point of information, nonenzymatic heating of the linear hydroxybilane in acidic solutions yields all four uroporphyrinogen isomers (I–IV) (numbered with regard to the side chain acidic groups), with the physiological uroporphyrinogen-III as the major isomer in a 1 : 2 : 4 : 1 ratio. Thus, if these molecules were forming in prebiotic media, the abundance of the uroporphyrinogen-III isomer may have aided in its subsequent biological processing on to heme, chlororophyll and vitamin B₁₂ for use by primordial organisms.

11.2.6 Uroporphyrinogen-III to Protoporphyrin IX

At this point organisms, *e.g.* humans, have arrived at the first tetrapyrrolic macrocycle

framework. On the one hand, this is a microbial biosynthetic branch point to siroheme and ultimately methanogen F430 9·10 (see Section 11.3). On the other hand, it is the “on pathway” precursor to protoporphyrin IX. On the route to that heme b type of macrocycle decoration, the next two enzymes remove six of the eight carboxylate groups as six CO₂. Four of these CO₂s arise as each acetate (carboxymethyl) side chain is decarboxylated to a methyl group.

At first blush this (the fifth) enzyme in the heme tetrapyrrole macrocyclic biosynthesis looks to have an easier job than the first four enzymes in the pathway. The decarboxylase acts on all four of the acetate (carboxymethyl) side chains to decarboxylate them, leaving the four propionic acid (carboxyethyl) side chains alone (Figure 11.11). The net result is conversion of an octacarboxy (octa-anionic) macrocycle to the tetra-anionic macrocycle known as coproporphyrinogen III (the prefix “copro” usually refers to dung or feces, reflecting the first isolation of this metabolite in porphyria variants 11). The four acetic acid side chains have become methyl groups, a net replacement of each carboxylate by a proton.

The final three enzymes in the heme biosynthetic pathway remove 6CO₂ and 6e⁻

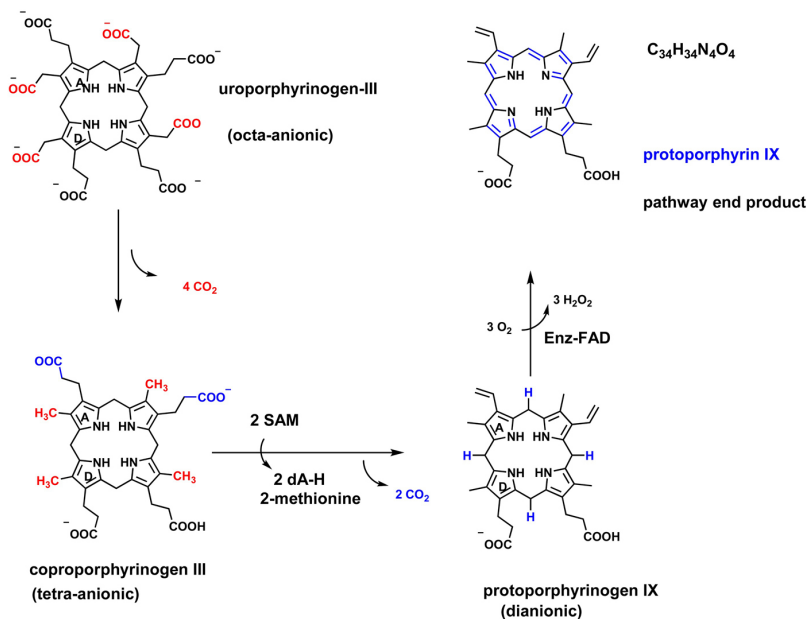


Figure 11.11 The final three enzymes in the heme biosynthetic pathway remove six CO₂ and six e⁻ from uroporphyrinogen III to arrive at protoporphyrin IX.

On further analysis the decarboxylation of an unactivated -CH₂COO⁻ to a methyl substituent is a reaction with a very high energy barrier. Lewis and Wolfenden extrapolated the slow nonenzymatic rate to a nonenzymatic half-life of 2.3 billion years. 12 The enzyme itself is on the slow end of catalysis at 10 turnovers per minute, but still boasts a catalytic:uncatalyzed ratio of 10²⁶. This puts it at the upper limit of catalytic proficiency for any enzyme known to date. How an enzyme evolves to improve from such a low uncatalyzed rate is not understood.

The presumption is that the enzyme can deliver a proton to bound substrate. While

the proton may rest mostly on the carboxylic oxygen, decarboxylation probably proceeds from the carboxylate so a very minor isomer protonated at C₂ (emphasizing the accessible carbanion resonance structure at C₂ of pyrrole rings) may be the species that actually decarboxylates. It is not known what surveillance mechanism the enzyme uses to ensure complete decarboxylation of all four of the acetate side chains while leaving the propionates intact.

Two of the four propionates, on the A and B rings, are next oxidatively decarboxylated to two vinyl groups by a radical *S*-adenosylmethionine (SAM)-dependent enzyme. The remaining step to protoporphyrin IX is flavoenzyme-mediated oxidation at the methylene bridges to create the 20 π electron, fully conjugated macrocycle **13** (Figure 11.12). Desaturation of the 14 π electron macrocycle to the 20 π electron tetrapyrrolic macrocycle protoporphyrin IX is thermodynamically favored (occurring as three 2e⁻ oxidative steps), creates a fully delocalized peripheral macrocyclic ring and gives the color–light absorbing properties of the final tetrapyrrolic ring metabolites as the pigments of life.

The FAD-dependent oxidase successively removes two electrons in three catalytic cycles to yield the heme pathway end product

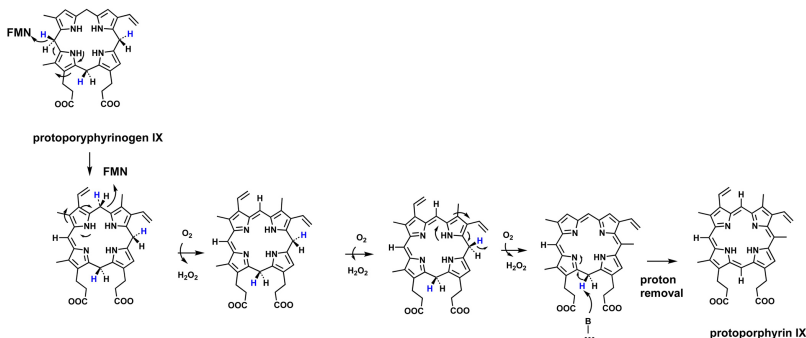


Figure 11.12 The terminal enzyme is an FAD-containing oxidase that reduces each of three molecules of O₂ by two electrons to H₂O₂ as three C=C double bonds are progressively introduced. Use of the pyrrole nitrogen lone pair in intramolecular assistance in hydride expulsion to N₅ of oxidized FAD reveals the underlying chemical logic for driving the macrocycle into fully conjugated protoporphyrin IX.

All told, 14 of the 48 carbons that are funneled into heme biosynthesis are expelled as CO₂, another underlying principle of this biosynthetic pathway. Of the eight starting glycine substrate molecules, four of the nitrogens are eliminated as four NH₃ in the hydroxybilane synthase reaction. The other four end up as the nitrogens in the four pyrrole rings.

11.3 Competing Chelataes: Insertion of Fe²⁺ or Mg²⁺

The protoporphyrin IX endpoint for heme biosynthesis is also a branch point for photosynthetic organisms that make chlorophyll. In plants that make both chlorophyll in chloroplasts and heme in mitochondria and cytoplasm, the partitioning is a competition between action of an iron chelatase vs. a magnesium chelatase (Figure 11.13A). We take up the ATP-dependent magnesium chelatase in Section 11.4.2 in the

further pathway logic to chlorophyll.

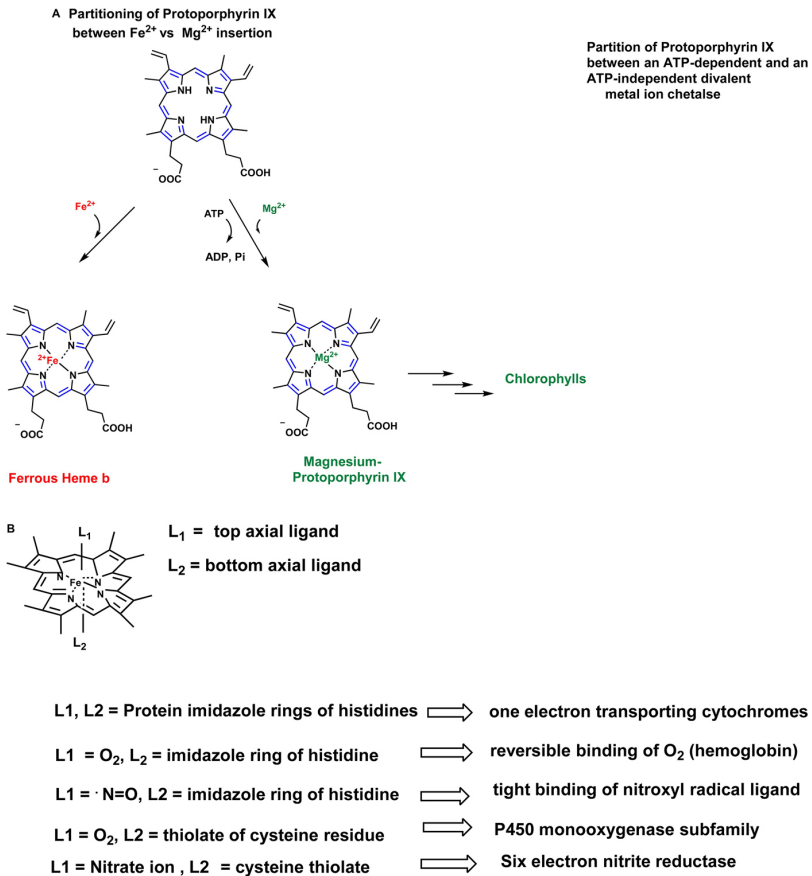


Figure 11.13 (A) Partitioning of “empty” protoporphyrin IX between cellular chelatases in plants that insert Fe^{2+} to make heme-b vs. magnesium chelatases that are ATP-dependent and vector the Mg^{2+} macrocycle on to chlorophylls. Humans have the iron chelatases but not the magnesium chelatases. (B) Five examples of variable top and bottom axial ligands to the iron in heme macrocycles, with attendant alteration in physiological redox functions.

The ferrochelatase 14 is independent of ATP, unlike the magnesium ion chelatase. Ferrochelatases across different kingdoms of life insert ferrous iron (Fe^{II}) into the equatorial plane of protoporphyrin IX to yield an embedded Fe^{II} – coordinated in the equatorial plane of the macrocycle via the four pyrrole nitrogens as ligands for iron (Figure 11.13A). For example, the human red cell ferrochelatase makes the ferrous heme cofactor for the hemoglobin that carries O_2 to all the tissues in the body. To insert Fe^{II} , the chelatase is thought to bind macrocyclic rings BCD in one plane and distort the A ring anywhere from 11° up to 36° out of plane, allowing insertion of the ferrous iron atom before the macrocyclic framework reverts to the low energy planar form and traps the central iron atom. There is evidence that the more abundant ferric ions in aerobic cells must first be reduced to air-sensitive Fe^{2+} , for which there is an initial nonheme binding site in ferrochelatase. This step helps select for Fe^{2+} insertion

into protoporphyrin IX over competing, abundant Zn^{2+} ions *in vivo*.

Once iron has been inserted, the resultant heme group binds to many different kinds of partner proteins, typically imparting the characteristic red-brown color to hemeproteins. The differential biologic functions of hemeproteins, as noted in Chapters 2–4 on inorganic nitrogen metabolism, can be modulated by modifications to the heme tetrapyrrolic macrocycle as in heme c, heme d1 (green colored) and P460 noted in Chapter 5.

Iron in both the 2+ and 3+ oxidation states in aqueous media prefers to be six-coordinate, with four equatorial ligands and top and bottom axial ligands, as does magnesium in the 2+ oxidation state. The individual heme proteins control heme reactivity indirectly by the cofactor binding protein microenvironments and more directly by the top and bottom axial ligands, 15 L1 and L2 in Figure 11.13B. When the top and bottom axial ligands are imidazole ring (nitrogens) side chains, the cellular functions of these cytochrome proteins are typically as one-electron transfer agents (e.g. b-type and c-type membrane bound cytochromes), between obligate one-electron donors and obligate one-electron acceptors. When the bottom axial ligand is a histidine imidazole side chain, but the top ligand is O_2 , bound reversibly, this is the ligand set for oxyhemoglobin, with O_2 dissociating as red cells encounter regions of low pO_2 in peripheral tissues. In general, O_2 binding to red cell hemoglobin and muscle myoglobin is reversible, with no redox chemistry. When one-electron transfer does occur mistakenly, superoxide ion and Fe^{3+} methemoglobin are formed. The Fe^{3+} must be reduced back to Fe^{2+} by methemoglobin reductase action for hemoglobin to regain functional status for O_2 binding.

When O_2 is the reversibly bound top axial ligand but the apoprotein provides a cysteine thiolate as the bottom axial ligand, this is the coordination geometry typical for the cytochrome P450 class of heme proteins that have catalytic activity for O_2 reduction, cleavage and transfer of an [OH] equivalent to substrates. We take up examples of hemeprotein *N*-oxygenases in the next chapter.

The fifth example of the L1, L2 inventory in a heme protein in Figure 11.13B is the heme catalytic group in nitrite reductase (see Chapter 4 for assimilatory nitrate reduction). The bottom axial ligand is a cysteine thiolate. The top axial ligand at the start of a catalytic cycle is a nitrite ion. It undergoes a series of six progressive one-electron reductions on the way to NH_3 . There are many other variants of heme proteins including human guanylate cyclases where the heme group is a signal receiver and the nitric oxide radical is the activator as the top axial ligand, discussed further in the next chapter also.

11.4 Two Biosynthetic Branchpoints Yield Tetrapyrrole Macrocycles with Equatorial Iron, Magnesium, Cobalt or Nickel

In addition to iron as a ligand equatorially situated between the four ring nitrogens in heme cofactors, chlorophyll A (Figure 11.1) with its equatorial magnesium cation (Mg^{2+}) may be the most abundant pigment on earth, given its central role in capture of light energy for photosynthesis by plants. The third pigment, vitamin B_{12} , is more specialized, made by microbes but remarkable for its redox-active cobalt (Co^{3+} in resting state). Most recently discovered is the nickel-containing tetrapyrrole macrocycle F430 that is the site of the last reductive step in methanogenesis by anaerobic bacterial methanogens (Figure 11.1). The biological chemistry available to the four tetrapyrrolic macrocycles is dramatically different, conditioned by the central metal cation. The central Fe, Co and Ni atoms are redox active. Mg^{2+} is redox inert in

biology but Mg-chlorophyll is the primary light absorbing center in photosynthesis. The four pigments absorb light of different wavelengths and so present as different colors. For example, chlorophyll a absorbs both red and blue light but not green light which it reflects. Hence, chlorophyll is green as are plants containing chlorophyll.

11.4.1 Branchpoint from Uroporphyrinogen-III to Siroheme, B₁₂, and F430

11.4.1.1 Siroheme

The first tetrapyrrolic biosynthetic intermediate in the heme biosynthetic pathway (Figure 11.14), uroporphyrinogen-III, is also the first branch point. 9:10:16:17 Uroporphyrinogen-III is a highly charged octacarboxy macrocyclic polyanion, compared with the dianionic porphyrin-IX and its metallated form heme b. Three enzymes take uroporphyrinogen to the siroheme cofactor for nitrite reductase (Figure 4.7).

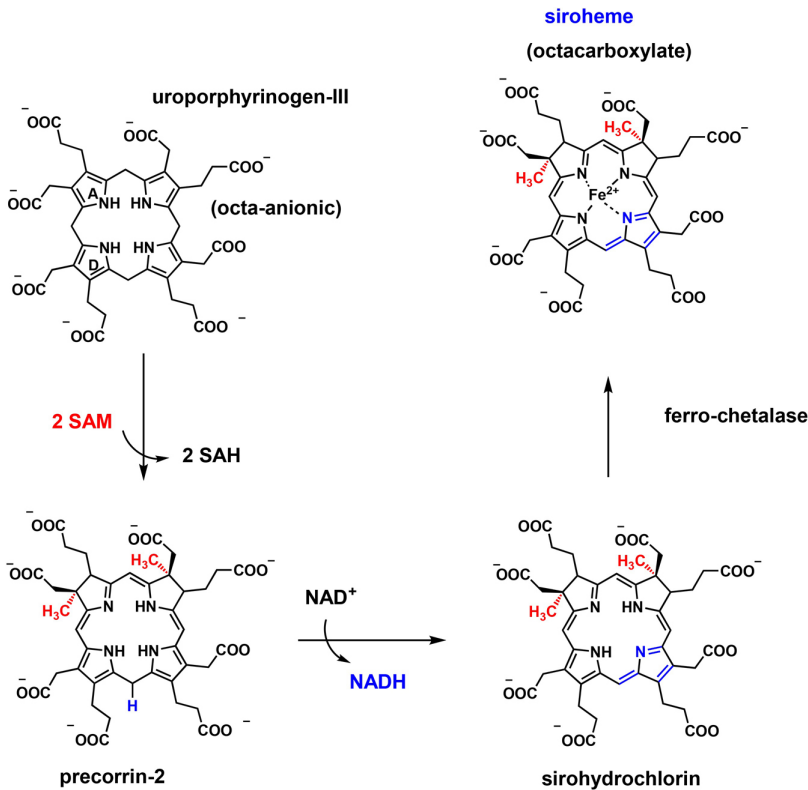


Figure 11.14 Partitioning at the level of uroporphyrinogen III between the heme pathway and routing some of the flux to siroheme biosynthesis Three enzymatic steps, two methylations and a hydride reduction, generate the tetrapyrrole macrocycle sirohydrochlorin. A specific iron chelatase inserts ferrous ion to create siroheme, the cofactor for nitrite reductase in the bacterial and plant nitrate assimilation pathways (Chapter 4).

The first enzyme catalyzes two radical-SAM methyltransferase steps that introduce methyl groups onto the sp^2 carbon of two of the four pyrroles, the A ring and B ring, rearranging the double bond configurations therein (Figure 11.14). This product is precorrin 2. The second enzyme is an NAD^+ -reducing dehydrogenase. The CH_2 methylene bridge between the C and D rings of precorrin 2 is oxidized to a CH group by removal of one of the two hydrogens as a hydride ion. The resultant macrocycle with extended conjugation is siroheme. The third step is conversion on to violet colored siroheme by a chelatase that inserts a ferrous iron atom into the equatorial plane (Figure 11.14). In some bacteria, such as *Salmonella enterica*, all three steps are catalyzed by one tripartite enzyme CysG. 16 Octa-anionic siroheme is the simplest tetrapyrrole macrocycle in terms of number of biosynthetic steps. It is the active site redox iron-macrocycle for six-electron nitrite reductases and sulfite reductases. It is also a late-stage intermediate for three other modified tetrapyrrole macrocycles.

11.4.1.2 Siroheme to Heme d1, Coenzyme F430 and Corrins

The uroporphyrinogen-III branch to precorrin 2 is also the route to several other modified tetrapyrrole macrocycles. That branchpoint metabolite (Figure 11.15), two steps away from siroheme, is likewise three steps away from the green (rather than red–brown) heme d1 18 in bacteria that practice inorganic nitrogen redox metabolism (Chapter 3).

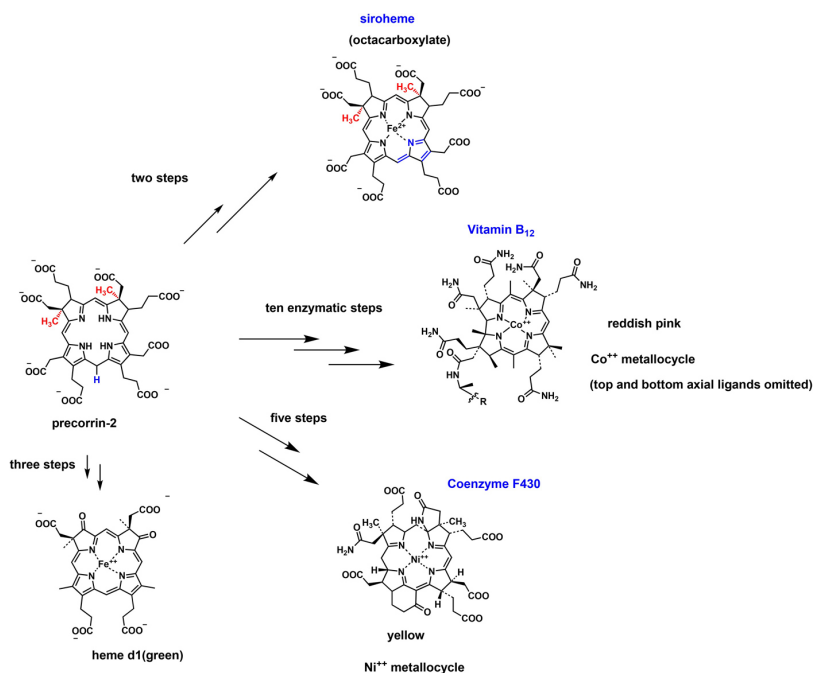


Figure 11.15 The siroheme diversion pathway also controls flux to three other tetrapyrrolic macrocyclic variants: cobalt-containing vitamin B₁₂ biosynthesis, nickel-containing F430 coenzyme in bacterial methanogenesis, or heme d1 in the nitrite reductase of denitrifying bacteria (Chapters 4 and 5).

Precorrin 2 is also five steps away from coenzyme F430 in methanogenic bacteria.

The path to F430 involves early metalation of precorrin 2 with a nickel²⁺ cation by a cation-specific chelatase as a way of drawing flux down the F430 pathway in methanogens. There is analogy in B₁₂ and F430 chemistry in that each of them undergoes enzymatic methylation to the organometallic methyl-B₁₂ and methyl-F430-organocobalt and organonickel tetrapyrroles, respectively.

The pathway from precorrin-2 to the mature corrin ring system of vitamin B₁₂ is much longer. There are ten enzymatic steps to introduce the large number of modifications, including the formal extrusion of the methene carbon between the A and D rings and the construction of the intramolecular dimethylbenzimidazole bottom axial ligand. Readers may consult a recent review on biosynthesis of modified tetrapyrrole macrocycles from the precorrin-2 intermediate stage for more insights and details. ⁹

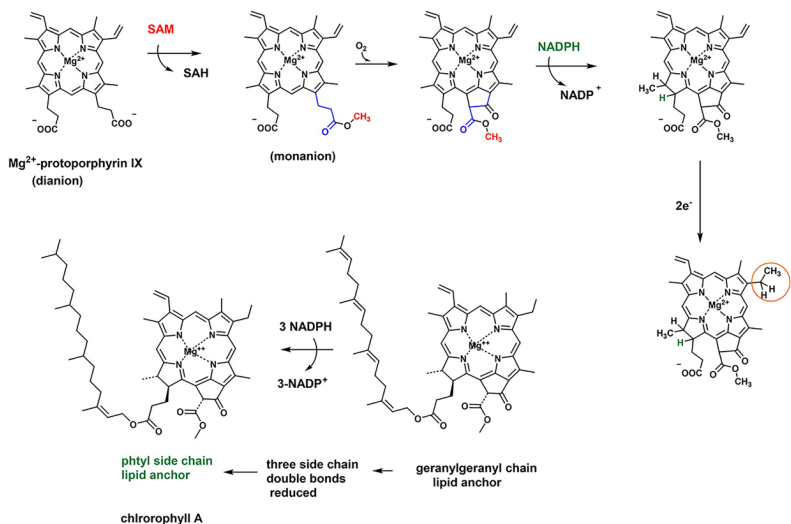
11.4.2 Branchpoint at Protoporphyrin IX to Chlorophylls

The second branch point occurs at the point of competition for metalation of protoporphyrin IX between ferrochelatase and magnesium ion chelatase. We have noted above the energy-independent insertion of an Fe^{II} ion into the equatorial plane of the macrocycle to generate heme.

In plant chloroplasts instead, a three-subunit Mg²⁺ chelatase competes for protoporphyrin IX (Figure 11.13A) and can draw it off to a Mg^{II} metalated porphyrin. ¹⁹ The Mg²⁺ chelatase, in contrast to the Fe²⁺ chelatase consumes several ATPs to effect successful insertion of the magnesium cation. The mechanism of coupling ATP hydrolysis to metal insertion is unclear but is thought to involve high energy, unstable conformers of the magnesium chelatase–porphyrin complex.

The Mg^{II}–protoporphyrin is then subjected to further maturation in plant chloroplasts (Figure 11.16A), including five reduction steps (four NADPHs and one reduced ferredoxin protein). First is a propionyl side chain carboxylate O-methylation, then oxidative cyclization of that propionyl moiety to a cyclopentanone ring fused to the lower right hand pyrrole ring, creating the characteristic green color of chlorophylls. ²⁰ Two successive reductions reduce the pyrrole C ring and the vinyl group substituent on the B ring.

A. Mg^{2+} -protoporphyrin IX to chlorophylls



B. chlorophyll A to chlorophyll B by consecutive oxygenase steps

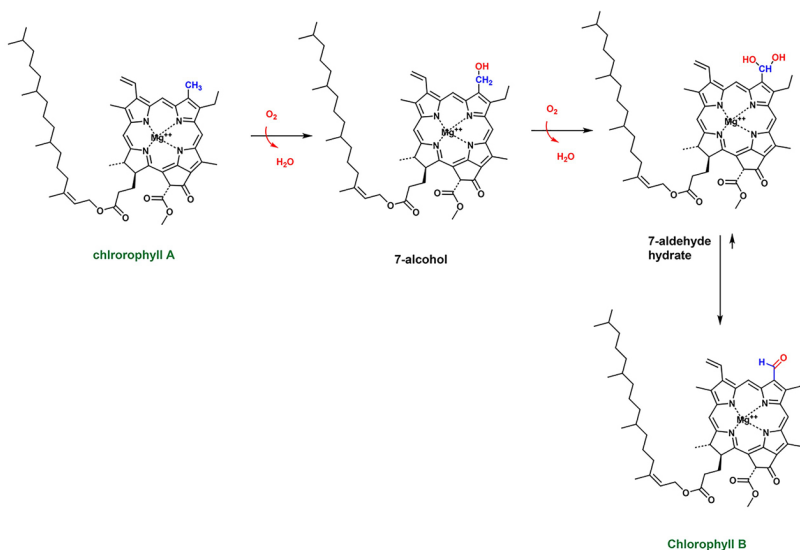


Figure 11.16 (A) Schematic of conversion of magnesium-chelated protoporphyrin-IX to chlorophyll A, involving methylation, oxygen-dependent cyclization of the C-ring propionate to a cyclopentanone, reduction of the D ring and of the vinyl group on the B ring before prenylation with the C₂₀ geranylgeranyl group that sets the membrane anchor. Triple reduction of the geranylgeranyl group to the saturated C₂₀ phytol chain improves stability of that membrane anchor to reactive oxygen species. (B). The second major chlorophyll compound in plant photosystems is chlorophyll B, available by double oxygenation of the C₇-CH₃ group in chlorophyll A to the 7-aldehyde in chlorophyll B, altering the visible absorption peak to allow chlorophyll B to absorb incident light at a lower wavelength maximum. Absorption peak spectra, reproduced from <https://>

The final two enzymatic steps in *chlorophyll A* formation are esterification with a C₂₀ geranylgeranyl side chain and then reduction of three of the four double bonds in that C₂₀ side chain *via* proton–hydride mechanisms involving three NADPH reductant molecules. This lipophilic tail anchors chlorophyll a into both of the two photosystems [PSI (approximately 60 chlorophylls) and PSII (30 chlorophylls)] in thylakoid membrane discs in plant chloroplasts. ²¹

Chlorophyll B differs from chlorophyll A by a double oxygenation of the methyl group on the B ring to give first the alcohol, and then the aldehyde substituent on that pyrrole (**Figure 11.16B**), shifting the wavelength maxima from 465 nm and 665 nm in chlorophyll A to 453 nm and 642 nm in chlorophyll B. Chlorophyll B is the subsidiary absorber of blue light while chlorophyll A absorbs red light and is the main donor of electrons to the photosynthetic electron transport chain.

The intimate details of light absorption and proton–electron excited state charge separation in photosynthesis are beyond the scope of this pyrrole-focused chapter. Interested readers can find extensive primary and review literature on topics from photophysics to chloroplast architecture in chlorophyll light harvesting functions.

11.5 Overview

During evolution the tetrapyrrolic macrocyclic framework has been diversified to serve many purposes – from light absorption in photosynthesis (chlorophyll), to the one-electron redox functions of heme macrocycles, to the six-electron reduction of inorganic sulfite and inorganic nitrite anions, to the organometallic intermediates in the contracted corrin macrocycle of vitamin B₁₂ and the methyl nickel intermediate in methanogenesis. The electronics of the macrocycles vary from a fully conjugated 20 π electron macrocyclic scaffold in hemes to the reduced macrocycles with partially interrupted conjugation in sirohemes, chlorophylls and F430 that tune the color and the electronics of the central Fe, Mg or Ni. The contraction of the corrin ring system is especially difficult biosynthetic chemistry to achieve and presumably sets an equatorial cavity customized to and necessary for enabling chemistry that utilizes three distinct redox states of the cobalt ion in the center of B₁₂ (see Chapter 10). Finally, we note that the early biosynthetic construction of the uroporphyrinogen-III framework yields an octa-anionic octacarboxy periphery to the tetrapyrrole macrocycle. The charge is modulated in two distinct modes. Six of the eight carboxylates are shed as CO₂ during heme and chlorophyll maturation. For the corrin ring system in vitamin B₁₂, on the other hand, enzymatic conversion of carboxylates to amides is an alternate route to a neutral scaffold.

The most noteworthy feature of tetrapyrrole macrocycle assembly is aminopyrrole chemistry, enabling the remarkable elongation of porphobilinogen to four tethered linear pyrrole units and the ring-closing macrocyclization step. In many ways aminopyrroles may be the most underappreciated of all the nitrogen heterocycles that mediate the chemistry of life. Their ability to function both as nucleophiles and electrophiles are at the center of the, often dazzling, chemical transformations: this is the carbanion chemistry enabled by aminopyrrole ring construction early in the pathway.

References

1. A. R. BattersbyTetrapyrroles: the pigments of life, *Nat. Prod. Rep.*, 2000, **17**, 507-526.
2. H. A. Dailey, T. A. Dailey, S. Gerdes, D. Jahn, M. Jahn, M. R. O'Brian and M. J. WarrenProkaryotic heme biosynthesis: multiple pathways to a common essential product, *Microbiol. Mol. Biol. Rev.*, 2017, **81**, e00048-16.
3. M. Petricek, K. Petrickova, L. Havlicek and J. FeisbergOccurrence of two 5-aminolevulinate biosynthetic pathways in *Streptomyces nodosus* subsp *asukaensis* is linked with the production of asukamycin, *J. Bacteriol.*, 2006, **188**, 5113-5123.
4. E. K. JaffeThe porphobilinogen synthase-catalyzed reaction mechanism, *Bioorg. Chem.*, 2004, **32**, 316-325.
5. N. Bung, A. Roy, B. Chen, D. Das, M. Pradhan, M. Yasuda, M. I. New, R. J. Desnick and G. BulsHuman hydroxybilane synthase: molecular dynamics of the pyrrole chain elongation identifies step-specific residues that cause AIP, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**(17), E4071-E4080.
6. R. Gill, S. Kolstoe, F. Mohammed, A. Al D Bass, J. E. Moseley, M. Sarwar, J. B. Cooper, S. P. Wood and P. M. Shoolingin-JordanStructure of human porphobilinogen deaminase at 2.8 angstroms: the molecular basis of acute intermittent porphyria, *Biochem. J.*, 2009, **420**, 17-25.
7. M. A. Matthews, H. Schubert, F. R. G. Whitby, K. J. Alexander, K. Schadick, H. A. Bergonia, J. D. Phillips and C. P. HillCrystal structure of human uroporphyrinogen III synthase, *EMBO J.*, 2001, **20**, 5832-5839.
8. S. D. Whatley, A. Roberts, D. H. Llewellyn, C. P. Bennett, C. Garrett and G. H. ElderNon-erythroid form of acute intermittent porphyria caused by promoter and frameshift mutations from the coding sequence of exon 1 of the HMBS gene, *Hum. Genet.*, 2000, **107**, 243-248.
9. D. A. Bryant, C. Hunter and M. J. WarrenBiosynthesis of the modified tetrapyrroles-the pigments of life, *J. Biol. Chem.*, 2020, **295**, 6888-6925.
10. S. J. Moore, S. Sowa, C. Schuchardt, E. Deery, A. D. Lawrence, J. V. Ramos, S. Billig, C. Birkemeyer, P. T. Chivers, M. J. Howard, S. E. Rigby, G. Layerb and M. J. WarrenElucidation of the biosynthesis of the methane catalyst F430, *Nature*, 2017, **543**, 78-82.
11. S. Sassa and A. KappasMolecular aspects of the inherited porphyrias, *J. Intern. Med.*, 2000, **247**, 169-178.
12. C. A. Lewis and R. WolfendenUroporphyrinogen decarboxylation as a benchmark for the catalytic proficiency of enzymes, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 17328-17333.
13. T. A. Dailey and H. A. DaileyIdentification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases and phytoene desaturase, *J. Biol. Chem.*, 1998, **273**, 13658-13662.
14. A. Medlkoek, L. Swartz, T. A. Dailey, H. A. Dailey and W. N. LanzilottaSubstrate interactions with human ferrochelatase, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1789-1793.
15. T. Li, H. L. Bonkovsky and J.-T. GuoStructural analysis of heme proteins: implications for design and prediction, *BMC Struct. Biol.*, 2011, **11**, 13.
16. J. M. Pennington, M. Kemp, L. McGarry, Y. Chen and E. StroupeSiroheme synthase orients substrates for dehydrogenase and chelatase activities in a common active site, *Nat. Commun.*, 2020, **11**(1), 864.
17. M. J. Warren, E. Roux, H. L. Schubert and J. C. Escalante-SemerenaThe biosynthesis of adenosylcobalamin (vitamin B12), *Nat. Prod. Rep.*, 2002, **19**, 390-412.
18. S. Bali, A. Lawrence, S. A. Lobo, L. M. Saraiva, B. T. Golding, D. J. Palmet, M. J. Howard, S. J. Ferguson and M. J. WarrenMolecular hijacking of siroheme for the synthesis of heme and d1 heme, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**,

19. N. B. P. Adams, A. Brindley, C. N. Hunter and J. D. Reid The catalytic power of magnesium chelatase: a benchmark for the AA⁺-ATPases, *FEBS Lett.*, 2016, **590**, 1687-1693.
20. C. Reinbothe, M. El Ballouri, F. Buhr, N. Murakia, J. Nomata, G. Kurisu, Y. Fujita and S. Reinbothe Chlorophyll biosynthesis; spotlight on protochlorophyllide reduction, *Trends Plant Sci.*, 2010, **15**, 614-621.
21. D. von Wettstein, S. Gough and C. G. Kannangara Chlorophyll biosynthesis, *Plant Cell*, 1995, **7**, 1039-1057.
22. I. Astner *et al.*, Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans, *EMBO J.*, 2005, **24**, 3166-3177.
23. D. Alexeev *et al.*, The Crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme, *J. Mol. Biol.*, 1988, **284**, 401-419.

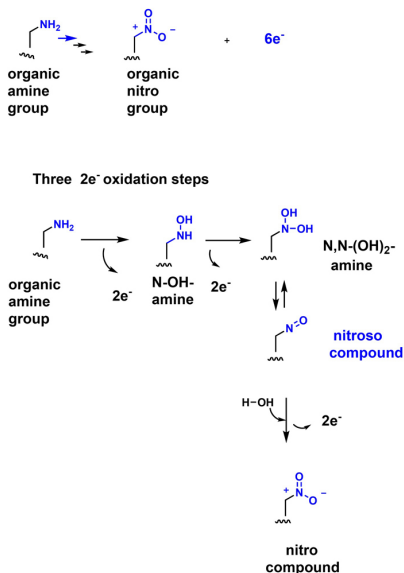
Organic Nitrogen Oxygenations

12.1 Iron-based vs. Flavin-based N-Oxygenases

All the organic amine, organic amide, organic amidine and guanidine enzymatic transformations described in the previous chapters, including the outlines of strategies for nucleic acid purine and pyrimidine base construction and the amide-bond forming chemistry of both proteinogenic and nonproteinogenic amino acid monomers are non-oxidative. The capture of those nucleophilic nitrogen groups by electrophilic fragments of cellular cosubstrates all happen at the $N = -3$ oxidation starting states (both primary amines and ammonia). In addition, the B vitamin heterocycles and the pyrrolic nitrogens in tetrapyrrole macrocycle biosynthetic pathways are rarely if ever involved in nitrogen atom oxidation or reduction.

However, just as inorganic nitrogen metabolism involves nitrogen at oxidation states that span eight electrons (NH_3 to NO_3^-), there are enzymes catalyzing six-electron (inter)conversions of organic nitrogen scaffolds, from amines RNH_2 ($N = -3$) to nitro compounds R-NO_2^- ($N = +3$) (Figure 12.1). In between are *N*-hydroxy, *N,N*-dihydroxy, and *N*-nitroso intermediate oxidation states (Figure 12.1). The cosubstrates are most often molecules of the ground state triplet O_2 , requiring enzymatic cofactors that can reductively activate molecular oxygen for one-electron transfer steps (Figure 12.1B). Also observed are oximes, nitrones and oxazolidinones with newly constructed N–O bonds through oxygenations (Figure 12.1).

A. six-electron oxidative traverse:



B. two common cofactors for N-oxygenases

three additional oxygenated N functional groups

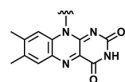
oxime



nitron



oxazolidinone



flavins
(organic coenzyme)



heme-iron
(cytochrome P450s)

Figure 12.1 (A) Some common *N*-oxygenated metabolites: *N*-hydroxylation, repeated to the *N,N*-dihydroxy metabolites with loss of water leads to *C*-nitroso compounds. Further two-electron oxidation yields *C*-nitro groups, six electrons oxidized from the starting amine. Additional *N*-oxy metabolites are oximes, nitrones and cyclic oxazolidinones. (B) The two most common cofactors for *N*-oxygenases are either flavin adenine dinucleotide (FAD) or the heme group in cytochrome P450 class monooxygenases.

We have seen in Chapters 2–5 on aerobic inorganic nitrogen redox metabolism that iron-based coenzymes, of both heme and nonheme architectures often serve in those electron transit roles. For the iron-based organic amine oxidases, enzymes most often use the cytochrome P450 subfamily of heme-based monooxygenases (see Figure 11.12) with a cysteine thiolate as the bottom axial ligand. An independently evolved set of iron oxygenases instead use nonheme binuclear iron clusters (Figure 12.2). Below, we will note bacterial *N*-hydroxylations that occur when they are in iron-limited environments. The *N*-hydroxy metabolites are sent out to scavenge and retrieve Fe^{II} ions to enable microbes, bacteria and fungi, to assemble the essential redox active heme and nonheme iron cofactors. 2

**typical cytochrome P450
ligation of bottom axial
cysteine thiolate**

**mononuclear
nonheme iron
center with O_2
bound**

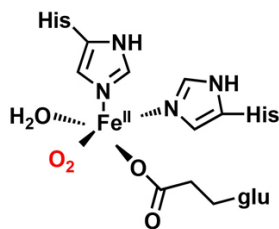
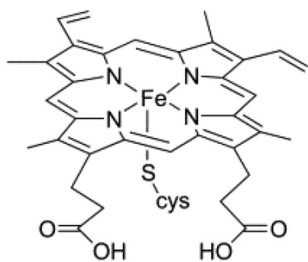


Figure 12.2 Two types of iron environments in iron-based oxygenases. Heme enzymes typically have a bottom axial thiolate ligand from a cysteine residue side chain with O_2 as the top axial ligand. Nonheme mononuclear iron enzymes typically have two histidine-imidazole and one aspartic acid or glutamic acid carboxylate ligands. A cosubstrate 2-oxoglutarate often occupies the fourth and fifth ligand positions, leaving the sixth site open for O_2 binding.

Additionally, many of the enzymes that catalyze *N*-hydroxylation of primary amines are metal free but instead substitute the redox active riboflavin (vitamin B₂)-based organic coenzyme forms (Figure 12.1B). ³ As discussed in Chapter 10, Section 10.3, these flavin cofactors can be reduced by NAD(P)H by two-electron paths. Then, the dihydroflavin coenzymes FADH₂ or FMNH₂ react with O_2 by one-electron paths, form covalent flavin-peroxy adducts, and transfer $[OH]^+$ equivalents to the nucleophilic substrate nitrogens, as noted in several examples later in this chapter.

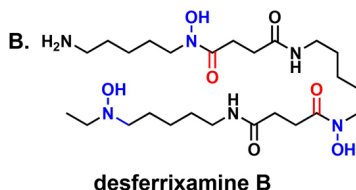
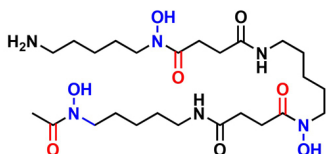
In parallel to inorganic nitrogen redox metabolism, many of the organic nitrogen oxygenation pathways occur in microbes, often as part of conditional, secondary metabolic pathways turned on in specific physiological contexts that confront certain bacterial species, such as scavenging for iron.

There are two major *N*-oxygenations that happen in human metabolism. One is the three-electron oxygenative fragmentation of L-arginine to citrulline and the nitric oxide radical. ⁴ We have seen the nitric oxide radical in three of the major redox pathways for inorganic nitrogen metabolites back in Chapters 3–5. Nitric oxide has a lifetime of approximately 5 seconds before dimerizing to N₂O, being captured by thiol groups or binding to heme iron centers and acting as a regulatory sensor molecule for lowering blood pressure (see later in this chapter). The second set of oxygenases are a family of five related human flavoprotein *N*-oxygenases that largely serve to detoxify xenobiotic molecules, ⁵ such as drugs, toxins and bacterial metabolites, by amine oxygenation. We take up these two enzymes after discussing the scope of microbial *N*-oxygenase actions.

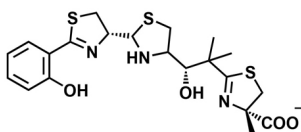
12.2 Two-electron Oxygenation of Amines to *N*-Hydroxylamines

Microbes need iron at internal levels of around one micromolar concentration to survive and grow. In times of low iron availability, bacteria and fungi assemble dedicated scaffolds of metabolites known as *siderophores* (from the Greek “bearers of iron”). The predominant form of iron in oxygenated aqueous environments is ferric ion (Fe^{3+}). The siderophores have evolved to capture ferric iron by multivalent chelation. Prominent among the siderophore chemical classes are: (1) *N*-hydroxy groups, often acylated to form bidentate hydroxamates, and (2) oxazoline and thiazoline heterocycles (Figure 12.3). (We have noted the oxazole and thiazole briefly in the chapters on nitrogen heterocycle biosynthesis.)

A. desferrioxamine B



yersiniabactin



vibriobactin

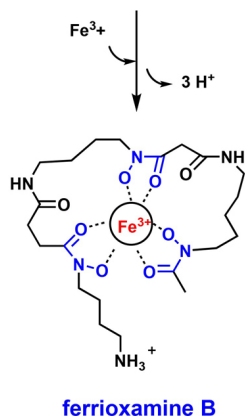
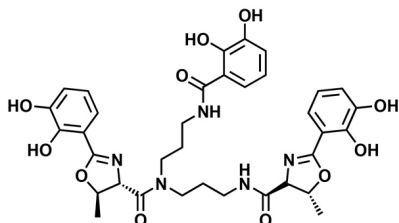


Figure 12.3 (A) Three types of siderophore structures that provide high affinity chelating ligands for ferric iron (the environmentally abundant form of dissolved iron): desferrioxamine B (hydroxamate group), yersiniabactin (salicylate, thiazoline) and vibriobactin (catechols). (B) Schematic of hexadentate coordination around the Fe^{III} ion by three hydroxamate groups in ferrioxamine B.

Two common building blocks for *N*-hydroxy siderophore frameworks are the amino acids L-ornithine and L-lysine. In the first steps of conditional siderophore biosynthetic pathways these two amino acids can be hydroxylated on N_5 or N_6 of their side chains, respectively (Figure 12.4). Three equivalents derived from N_6 -OH lysine 6 are found in the siderophore desferrioxamine B 7 (from *Streptomyces pilosus*) (Figure 12.3). The N_6 -OH-lysines have undergone pyridoxal phosphate (PLP)-enzyme mediated decarboxylation and then *N*-acylation by one acetyl CoA and two succinyl CoA units, resulting in three internal bidentate hydroxamate units in desferrioxamine B. The trio of hydroxamates can fold around an Fe^{III} ion to fill its six-coordinate ligand sphere

(Figure 12.3B) with sufficient affinity to scavenge iron even at 10^{-15} M external levels.

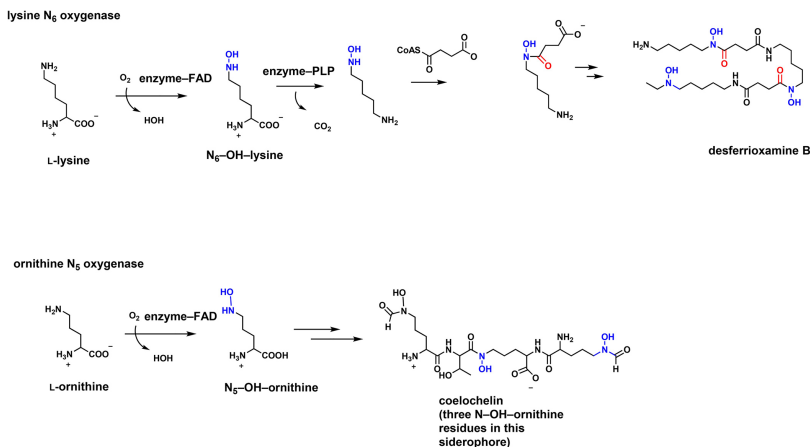


Figure 12.4 Lysine N₆ oxygenase action in desferrioxamine biosynthesis; ornithine N₅ oxygenase participates in biosynthesis of the iron scavenging coelochelin from *Streptomyces coelicolor* by providing three equivalents of N₅-OH-ornithine.

Correspondingly, enzymatic ornithine hydroxylation at N₅ occurs in the soil bacteria *Pseudomonas aeruginosa* ⁸ and *Streptomyces coelicolor* ⁹. Three such N-hydroxy amino acid metabolites are combined with an intervening threonine residue to yield the siderophore coelochelin. Two of the three N-OH-ornithine residues have been also N-formylated (via N-formyl-THF, Chapter 10) (Figure 12.4).

12.2.1 An Unusual N–O–Sugar Linkage in the Ene-diyne Antitumor Antibiotic Calicheamicin

The class of enediyne antitumor antibiotics are among the most potent natural cytotoxic agents known (Figure 12.5A). Their target is DNA, for both single- and double-strand breaks by abstraction of hydrogen atoms as the ene-diyne moieties undergo diradical aromatizing cyclizations (Figure 12.5), studied abiotically by Bergman and colleagues. ¹⁰ The ene-diyne chemistry and pharmacology topics are beyond the scope of this volume. ¹¹ However, one feature of calicheamicin, among many that are noteworthy, is an unusual –N–O– linkage between two of the sugars in the backbone. Figure 12.5B indicates that an unusual N-hydroxylation of a C₄-aminoglucose moiety occurs at some point in the biosynthetic assembly. ¹² At a later stage the oxygen of that hydroxylamine grouping acts as nucleophile towards a deoxyhexose unit electrophilic at C₁. The result is a C₁–O–N–C₄ unusual disaccharide part structure in calicheamicin Figure 12.5B. (Equally unusual is the 4-thio-linkage in that sugar moiety.)

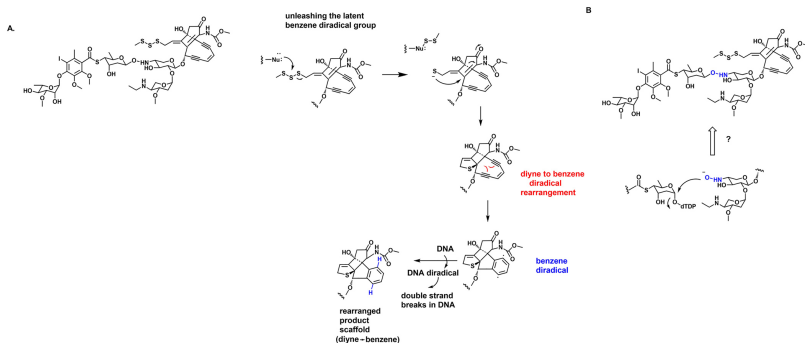


Figure 12.5 (A) Calicheamycin is a member of the ene-diyne class of antibiotics. These DNA-targeting agents are latent generators of 1,4-benzene diradicals. In turn, that radical pair abstracts hydrogen atoms from adjacent DNA bases, ultimately inducing single and double strand breaks by radical chemistry. (B) An unusual N–O linkage between two hexose residues in calicheamycin might arise by an N–OH sugar acting as a nucleophile.

12.2.2 Lysine to *N*-Hydroxy-pipecolate in Plant Defense

When plants sense attack by predators, *e.g.* feeding insects, they respond with a variety of chemical defenses. Some of them are newly synthesized phytoalexins that signal to other parts of the same plant or to neighboring plants. Some are preformed scaffolds termed phytoanticipins that are released essentially instantaneously ([Figure 12.6](#)). Some are precursors to reactive molecules to chemically ward off the predators. Glucosinolates, discussed below, fall into this category of nitrogenous defense agents [13](#) ([Figure 12.6](#)).

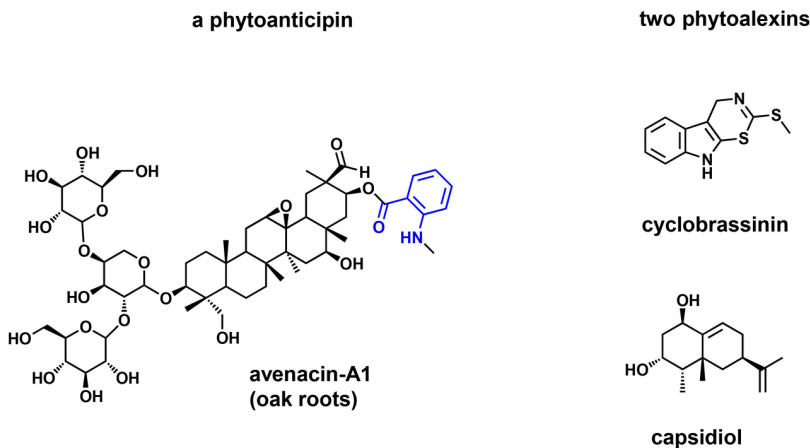


Figure 12.6 Representative phytoanticipin and phytoalexins produced in response to plant sensing of predations. Phytoanticipins are made and stored in the plants to be released immediately. Phytoalexin genes are turned on when predators are sensed.

Recently, the simple six-carbon cyclic amino acid pipecolate has been shown to serve

as a substrate for enzymatic *N*-hydroxylation (Figure 12.7). The resultant *N*-OH-pipecolate is a systemic signaling molecule for plants to turn on sets of defensive genes. 14 The two-step conversion of the common metabolite *L*-lysine to *N*-hydroxy-*L*-pipecolate illustrates two iterative aspects of the N_6 - NH_2 group of lysine as a nucleophile, first towards carbon and second towards oxygen.

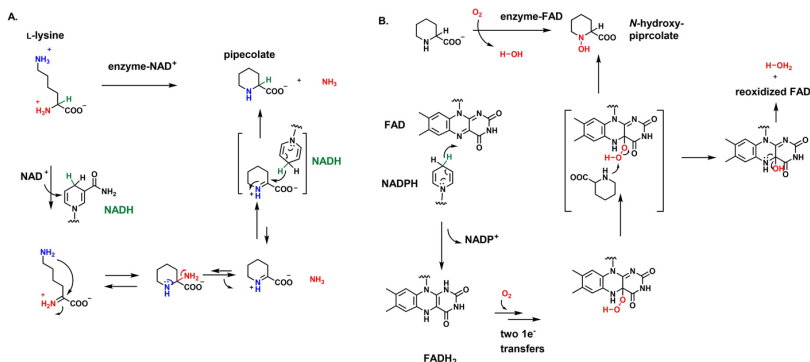


Figure 12.7 (A) Lysine to pipecolate with retention of N_6 nitrogen; (B) *N*-hydroxylation of pipecolate by a FAD-dependent oxygenase in plant metabolism.

The pathway starts with enzymatic conversion of lysine to pipecolate, the next highest homolog of the more familiar cyclic amino acid proline 15 (Figure 12.7A). As depicted, the N_6 - NH_2 of lysine remains in pipecolate while the N_2 - NH_2 has been ejected as NH_3 . The chemical logic to achieve C_2 - NH_2 bond labilization is to oxidize the C_2 -amine to an imine linkage. This occurs by enzyme-catalyzed transfer of the C_2 -H as a hydride ion directly to bound NAD^+ , creating the bound acyclic imine and $NADH$. Now, the deprotonated form of the N_6 side chain amine can act as a nucleophile at the newly formed $C_2=N$ imine to create the indicated tetrahedral adduct. That adduct can decompose either backwards or forwards. In the forward direction, the C_2 - NH_2 is ejected as released NH_3 . (Ejection of the C_6 - NH_2 constitutes the backwards decomposition.) This action creates the second iminium ion intermediate, this time as the cyclic imine. Hydride transfer back from $NADH$ to the electrophilic C_2 reduces that imine with chiral control to *L*-pipecolate.

The second enzyme in the pathway is a flavin-dependent monooxygenase. (It also uses an $NADH$ equivalent but this time it gets consumed as a stoichiometric reductant, rather than being regenerated as in the pipecolate synthase reaction.) $NADH$ reduces enzyme bound flavin adenine dinucleotide (FAD) to create the O_2 -reactive dihydroflavin oxidation state. A pair of one-electron transfers yields the covalent FAD -OOH adduct which then delivers a nascent $[OH]^+$ equivalent to the nucleophilic nitrogen of pipecolate 14 (Figure 12.7B). Thus, in two enzymatic steps, lysine has been converted to a cyclic amino acid with the C_2 - NH_2 labilized as NH_3 while N_6 is incorporated into the pipecolate scaffold. The original N_6 of *L*-lysine has acted iteratively as a nucleophile towards two kinds of cellular electrophiles: once towards the C_2 -imine, then once towards the FAD -OOH adduct.

12.3 Six-electron Oxidations of Amines to Nitroso and Nitro Groups

Strains of *Burkholderia thailandensis* correspondingly elaborate and export the siderophore family of malleobactins 16,17 (Figure 12.8), built from two *N*-5-hydroxyornithine residues via a nonribosomal tetrapeptide assembly line that also activates one β -hydroxyaspartate and one serine residue. The chain termination step for the nonribosomal peptide synthetase (NRPS) assembly line involves amide bond formation with 1,4-diaminobutane as a nucleophile (derived from ornithine enzymatic decarboxylation). The second *N*-OH ornithine residue, but not the first, in the initial siderophore product malleobactin B has been *N*-formylated to create a bidentate hydroxamate ligand for Fe^{II} binding. The *N*-terminal $\text{N}_5\text{-OH-Orn}$ residue has instead been acetylated at nitrogen₂.

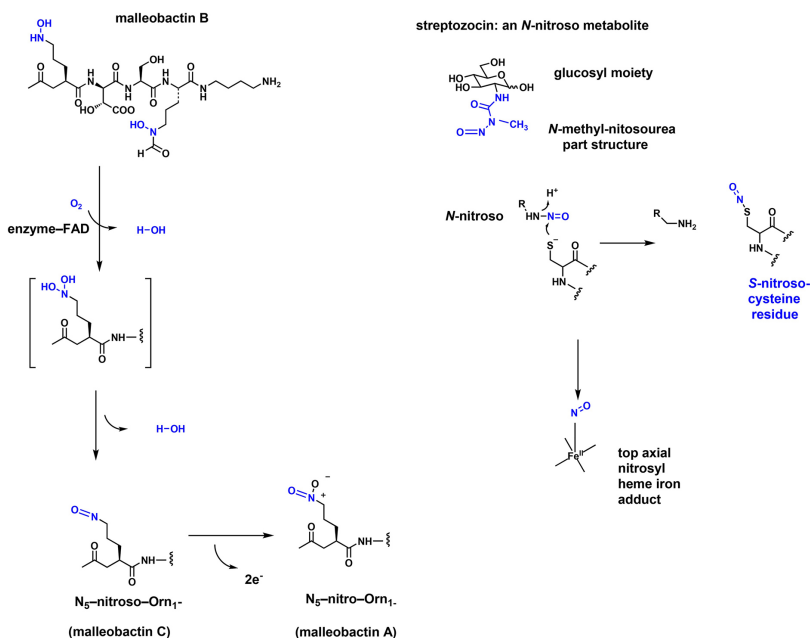


Figure 12.8 Conversion of malleobactin B to malleobactin A involves oxidation of an $\text{N}_5\text{-OH-Orn}_1$ residue to a nitro-ornithine residue by way of a C-nitroso intermediate. Streptozocin is an *N*-nitroso natural product (embedded in an *N*-methyl nitrosyl urea group) with nitrosyl transfer reactivity.

The $\text{N}_5\text{-OH-Orn}_1$ residue is subjected to further *N*-oxygenation, generating transiently the *N,N*-dihydroxy adduct, at the expense of reductive activation and splitting of another molecule of O_2 . The FAD-containing ornithine *N*-oxygenase that carries out the first *N*-oxygenation will likewise perform this second *N*-oxygenation when malleobactin B accumulates. Most *N,N*-dihydroxy species undergo facile nonenzymatic loss of water to yield the *nitroso* species and that is the case for malleobactin C.

The nitroso linkage in malleobactin C is a C-nitroso linkage. One can also find N-N=O linkages in microbial metabolites. That is the case in streptozocin, an actinomycete metabolite that has been used for treatment of pancreatic tumors 18 (Figure 12.8). The *N*-nitroso group is the reactive nitrosylating agent. It can be captured as an electrophile by protein thiol groups or alcohols to S-nitroso or O-nitrite ester groups, respectively (Figure 12.8). Analogously, the nitroso group can react with metals as in tetrapyrrole macrocycles to yield axial nitrosyl complexes that can

interrupt the cellular functions of the affected heme proteins.

The third oxygenation event in the malleobactin pathway leads to malleobactin A (Figure 12.8). The nitroso group has been further oxidized by two electrons to the N_5 -nitro-Orn₁ residue. Malleobactin A, the end product of this siderophore biosynthetic pathway, has one N_5 -nitro-Orn and one N_5 -N-OH, N -formyl-Orn residue. In terms of oxidizing steps, the conversion of Orn to N_5 -OH-Orn is a two-electron oxidation ($N = -1$ oxidation state). The second N -hydroxylation to the N,N -dihydroxy nascent product, and its dehydration to the C -nitroso-Orn₁, is a further two-electron oxidation. The oxidation state of the nitroso nitrogen = +1 (Figure 12.8). The final conversion of C -nitroso to C -nitro constitutes the third two-electron oxidation of N_5 of Orn₁. The oxidation state of a nitro group nitrogen is +3. The total six-electron traverse from the amino C -NH₂ to the C -NO₂ – takes nitrogen from the -3 oxidation state to the $+3$ oxidation state, whence the nitrogen is electrophilic not nucleophilic.

While nitro groups are rare in multicellular organisms, over 200 nitro-containing metabolite scaffolds are known from microbial metabolism, ¹ including the malleobactin A siderophore noted above. One of the long-standing examples is in the antibiotic chloramphenicol (Figure 12.9). The nitrophenyl group is a typical nitroaromatic moiety.

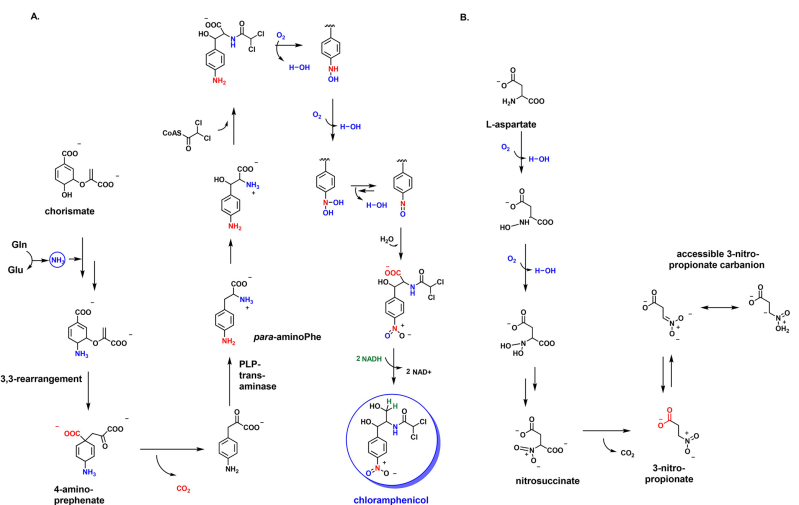


Figure 12.9 Two examples of amine oxidations to nitro groups. (A) Biosynthesis of the antibiotic chloramphenicol involves amination of chorismate, a 3,3-Claisen rearrangement, and aromatizing decarboxylation. Transamination and beta hydroxylation precedes the six-electron oxidation of the *para*-amino to the *para*-nitro group. A double reduction of the carboxylate to the alcohol yields the protein synthesis chain-terminating chloramphenicol. (B) Oxidation of the amino group *via* double oxygenation leads to the nitroso and then the nitro group of nitrosuccinate. Decarboxylation produces nitropropionate where the C₃ carbanion is stabilized by resonance with the *aci* nitro form.

It arises from the secondary metabolite *para*-amino-phenylalanine. In turn *para*-amino-phenylalanine arises from the same 4-amino-4-deoxychorismate that is the precursor to *para*-aminobenzoic acid (PABA). The divergence occurs when 4-amino-deoxy chorismate eliminates pyruvate while aromatizing to PABA vs. undergoing the directed 3,3-Claisen rearrangement that leads to amino-prephenate (Figure 12.9). Decarboxylative aromatization yields *para*-amino-phenylpyruvate, a transaminase-

mediated reductive amination away from *para*-amino-phenylalanine.

The two amino groups in *para*-amino-phenylalanine undergo two distinct fates. The C₂-amine acts as a nucleophile towards dichloroacetyl CoA to form the dichloroacetyl amide linkage in chloramphenicol. Orthogonally, the *para*-amino group is successively oxygenated three times to the N-OH, *N,N*, dihydroxy = nitroso, to the final nitrate anion. ¹⁹ The iron based oxygenase is neither a heme type nor a mononuclear iron oxygenase. Instead, it contains a bridged diiron cofactor to carry out the three oxygenative steps. ²⁰ The final two steps in formation of the chloramphenicol antibiotic scaffold are a double reduction of the carboxylate to the aldehyde and then to the primary alcohol.

A second example in Figure 12.9 is the fungal production of 3-nitropropionate, an antimetabolite of succinate. ²¹ The pathway starts with the metabolically central amino acid L-aspartate and proceeds by three successive two-electron oxidation-oxygenation steps to nitrosuccinate. The nitrosuccinate is then subjected to enzymatic decarboxylation to 3-nitropropionate and CO₂. Nitropropionate can tautomerize to the *aci*-carboxylate form that has a resonance contributor as the C₃ carbanion, an unusually accessible carbanion.

12.4 Nitrones and Oximes

Nitrones and oximes are two additional N-O linkage types that are installed biosynthetically in certain classes of nitrogen-containing natural products. ²² Figure 12.10 illustrates that nitrone groups are four-electron oxidized relative to the starting amines and presumably devolve from the kind of *N,N*-dihydroxyamine species noted above. Nitrones by definition are a >C=N-O⁻ functional group, the *N*-oxide of an imine. Oximes constitute a related C=N-OH grouping, formally arising from an aldehyde or ketone and hydroxylamine (Figure 12.10A). Neither nitrogen can be directly oxygenated because they are non-nucleophilic. They must come from amine precursors.

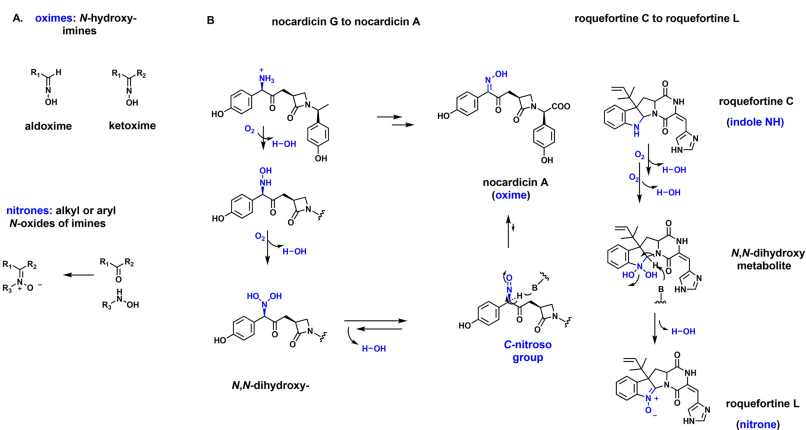


Figure 12.10 (A) Structures of oximes and nitrones. (B, left) Conversion of nocardicin G to nocardicin A involves two tandem oxygenations of the amino group of the phenylglycyl moiety to the C-nitroso group followed by rearrangement to the oxime. (B, right) The conversion of roquefortine C to roquefortine L involves tandem oxygenation of the indole NH to the *N,N*-dihydroxy metabolite that loses water across the N₁-C₂ bond to form the nitrone.

Two examples are (1) the formation of the oxime linkage in the antibiotic nocardicin A from the amine grouping in nocardicin G [23](#) and (2) the conversion of the amine in roquefortine C to the nitron linkage in roquefortine L [24](#) ([Figure 12.10B](#)).

The nonribosomally generated nocardicin has the now classic β -lactam as the chemical warhead for disruption of bacterial cell walls. It also has an unusual phenylalanyl moiety as the *N*-terminus with its free amino group. Enzymatic *N*-hydroxylation happens iteratively to give the *N,N*-dihydroxy metabolite from which loss of water yields the indicated nitroso structure ([Figure 12.10B](#)). Conversion of the four-electron-oxidized nitroso to the oxime structure is pictured as an isomerization by C–H bond cleavage to convert the CH–N=O nitroso moiety to the C=N–OH oxime group. [23](#) This scheme shows the intimate interrelationship between C-nitroso and oxime groups.

Roquefort is among the most famous of the French blue cheeses, restricted in appellation to particular regions of southwestern France. The mold *Penicillium roqueforti* is a main flavor generator in these cheeses. Minor mold metabolites can also be toxic as is the case for the family of nonribosomally generated diketopiperazines of the roquefortine family ([Figure 12.10B](#)). Roquefortine C is a diketopiperazine between tryptophan and histidine (see [Figure 9.8](#) for tryptophan–X diketopiperazines). The C₃ position of the tryptophan side chain has undergone a prenylation with concomitant attack by one of the diketopiperazine nitrogens on C₂ of the indole ring of tryptophan. Roquefortine C can be converted on to a series of oxygenated roquefortine metabolites, including roquefortine L [24](#) ([Figure 12.10B](#)). The N₁ nitrogen of the pyrrolidine ring acts as nucleophile twice in a row to give the *N*-hydroxy and then the nascent *N,N*-dihydroxy product. Loss of the elements of water from C₂ and N₁ generate the C₂–N₁ endocyclic iminium group with one oxyanion still attached to N₁. This is roquefortine L with a nitron moiety as the remnant of the *N,N*-dihydroxy immediate product.

12.5 Glucosinolates

Related C=N–O– linkages are found in a family of plant amino acid-derived defensive molecules known as glucosinolates because of the presence of a thioglucosyl moiety. [13·25](#) The active electrophile is a *thiohydroxamic acid* group, trapped between two protecting moieties: a glucosyl group attached to the thiol sulfur and a sulfate group attached to the oxime oxygen ([Figure 12.11A](#)). The structure of a tryptophan-derived glucosinolate is shown at the bottom of [Figure 12.11A](#). Hydrolytic cleavage of the thioglucoside bond, when plants such as broccoli, kale, and brussels sprouts are under attack, releases the latent warhead that unravels to potent electrophilic isothiocyanate groups ([Figure 12.11B](#)), among other fragments.

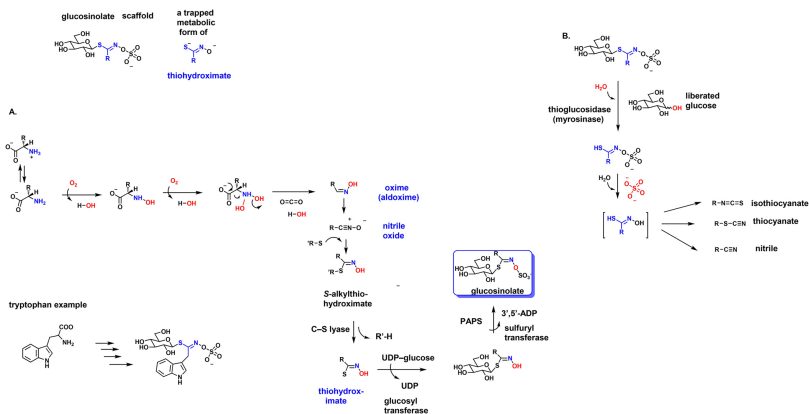


Figure 12.11 Glucosinolate formation and subsequent hydrolytic release of the thiohydroximate electrophile. (A) Double *N*-hydroxylation of amino acids to the *N,N*-dihydroxy metabolite sets up decarboxylation and loss of water to the aldoxime. Oxidation to the nitrile oxide creates a molecule with an electrophilic carbon. The carbon electrophile is capturable, by the thiolate of glutathione. Subsequent C–S lyase action (PLP-dependent) frees up the thiolate to act a second time as nucleophile, this time towards UDP–glucose to create the β -thioglycoside. Finally, for good measure *O*-sulfurylation produces the doubly protected glucosinolate. (B) Glucosinolate activations occur in two enzymatic steps. First the thioglycoside link is hydrolyzed enzymatically. Then a sulfatase acts to liberate the free thiohydroximate. That species can break down to a number of products, including the isothiocyanates that are electrophilic at the central carbon.

Biosynthesis of stored glucosinolates [26](#) starts, for example, with tryptophan double *N*-hydroxylation ([Figure 12.11A](#)), loss of water to the *C*₂-nitroso compound and catalyzed isomerization to the oxime. Further oxidation by a plant cytochrome P450 gives the unusual nitrile oxide, capturable on the nitrile carbon by the thiolate anion of glutathione, setting the S–C bond in the resultant *S*-alkyl thiohydroximate adduct. We take up the mechanism of these three cytochrome P450-mediated *N*-oxygenations in Chapters 13 and 14. Successive peptidase–isopeptidase action on the tripeptide glutathionyl moiety leads to the thiohydroximate.

Two steps of protective group additions, one on sulfur, one on oxygen, are catalyzed by successive enzymes. The thiol is still nucleophilic and can capture the glucosyl moiety at *C*₁ of UDP-glucose, creating the protected *S*-glycoside. The second protective step is to transfer a sulfuryl group from 3′-phospho-5′-adenosinephosphosulfate (PAPS, see Chapter 2) to the thiohydroxime oxygen and arrive at the mature glucosinolate. Both sulfur atoms (the thiol glucoside and the *O*-sulfate ester) are protecting groups keeping the intrinsic isothiocyanate warhead doubly latent. The coalescence of nitrogen, oxygen and sulfur enzyme chemistry has led to the unusual thiohydroximate warhead.

The uncovering of the glucosinolate warhead begins [27](#) when the thioglucosidase myrosinase is allowed to reach the glucosinolate scaffold as a membrane separating the two compartments is breached in plant alarm mode ([Figure 12.11B](#)). The free thiol is thereby liberated. Subsequent sulfatase action on the *O*-sulfate ester uncovers the rest of the thiohydroximate. As shown, that can decompose, directed in part by chaperone proteins in particular brassica plant species to isothiocyanates, thiocyanate, nitriles and other minor products (not shown). In particular, the isothiocyanates are reactive electrophiles that can capture both thiol and amine groups of nearby proteins, *e.g.* on predator surfaces, acting as pronounced antifeedants.

12.6 Ornithine and Lysine to Hydrazine Derivatives via N–O Linkages

Some N–O functional groups get converted to N–N functional groups. Authoritative and comprehensive reviews have recently appeared on biological routes to hydrazide, diazo, azoxy, hydrazine and *N*-nitroso functional groups [22](#) (Figure 12.12). Interested readers are referred in particular to the 2018 review by Waldman *et al.* [22](#) Here, we note only two examples during which the ornithine–N₅ and lysine–N₆ side chain *N*-hydroxylations are preambles to formation of N–N hydrazine groups. (In Chapter 4 we discussed the enzymatic conversion of hydroxylamine to hydrazine, an inorganic N–O to N–N change.)

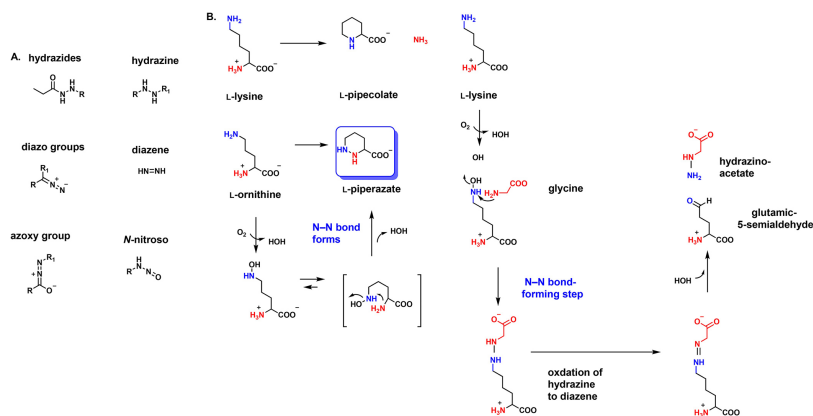


Figure 12.12 Some N–N groups found in microbial metabolites: (A) azoxy, hydrazide, diazo, hydrazine, diazene, *N*-nitroso groups. (B) Proposed formation of piperazate from *N*-hydroxy-ornithine.

In the biosynthesis of the nonribosomal antifungal hexapeptide kutzneride [28](#) the *Kutzneria* soil bacteria assemble that scaffold from all six nonproteinogenic amino acid monomers, one of which is piperazic acid (Figure 12.12B). Piperazate is the hydrazino analog of pipecolate whose biosynthesis from lysine is outlined in Figure 12.7. Instead, for the hydrazine linkage, L-ornithine is the starting substrate for enzymatic cyclization to piperazate. Both nitrogens are retained in a new N–N hydrazine linkage. (Recall that free hydrazine is a proposed intermediate both in nitrogenase action and in the anammox pathway described in Chapter 3.) L-Ornithine is oxygenated at N₅ by a paradigmatic flavoenzyme *N*-monooxygenase, KtzI. The *N*-hydroxy-ornithine is then the precursor to L-piperazate. The conversion is mediated by a heme protein KtzT. [29,30](#) It may be that the N₅–OH group of the N–OH–Orn becomes an axial ligand to the heme iron and is thereby polarized for N₅–O bond cleavage on intramolecular attack by the N₂ amine nitrogen (in analogy to the mechanism for conversion of NH₂OH to NH₂NH₂ described in Chapter 4).

Correspondingly, L-lysine is *N*-oxygenated at N₆ on the way to the microbial metabolite hydrazinoacetate [31](#) (Figure 12.12B). The next enzyme uses glycine as a nucleophilic cosubstrate in a proposed attack on the N₆ of N₆–OH-lysine. The resultant diacid has a hydrazine linkage. Oxidative conversion of the C₆–N–N-linkage to a C–N=N– group would be the preamble to hydrolysis to the observed products amino adipic semialdehyde and hydrazinoacetate. Note that in both examples the –N–

OH nitrogen is proposed to react as an electrophile not a nucleophile with cleavage of the N–OH bond, a reversal of nitrogen reactivity resulting from initial oxidation to the N–OH group.

12.7 Arginine N-Hydroxylation: to D-Cycloserine

The side chain guanidinium group of arginine with its pKa value of 12.5 is 99.999% protonated at pH 7.5. The free base form is so rare under physiological conditions that it is not available for significant chemistry as a nucleophile. Yet, we have noted in Chapter 8 that arginine kinases can, in effect, fleetingly create a deprotonated guanidine equivalent such that the terminal nitrogen attacks P γ of ATP in the active site of arginine kinases to produce the phosphagen arginine–N–PO₃^{2–} (Figure 8.8).

Analogously, one would not *a priori* feature the arginine guanidinium side chain as a nitrogen nucleophile for *N*-hydroxylation. Nonetheless, two enzymatic examples, one microbial, the other mammalian and human, reveal arginine side chain *N*-hydroxylation. In each case the *N*-hydroxy species is the prelude to two quite distinct products: an oxazolidinone in D-cycloserine vs. the odd-electron free radical nitric oxide (Figure 12.13).

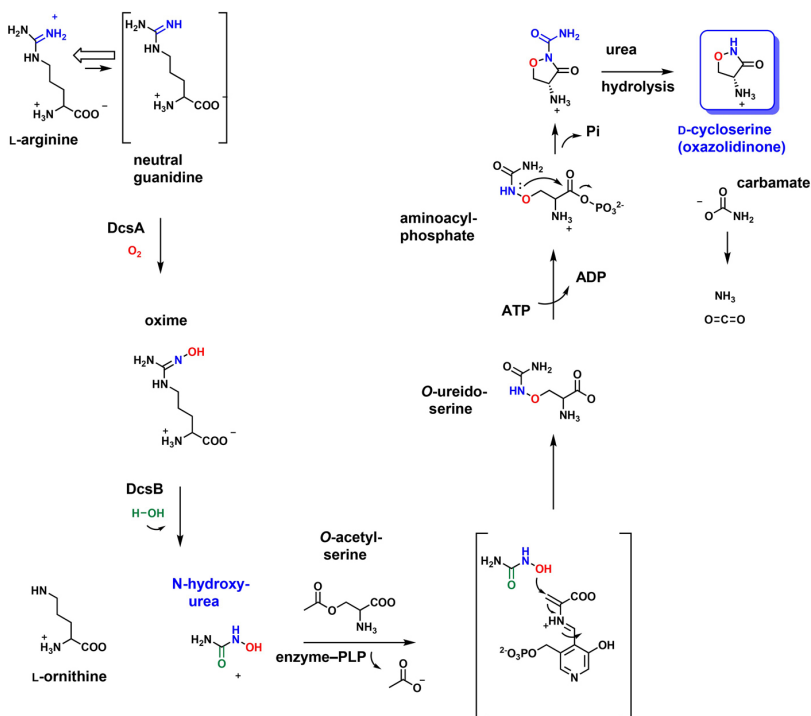


Figure 12.13 Conversion of arginine to the antibiotic D-cycloserine: *N*-hydroxylation followed by hydrolysis gives *N*-hydroxyurea that then acts as a nucleophile to form ureido serine. ATP promotes cyclization of the aminoacyl phosphate to the amido derivative of cycloserine (as part of the ureido group). Hydrolysis of the ureido group gives D-cycloserine and carbamate.

Streptomycete bacterial strains make the cyclic amino-oxazolidinone known as D-cycloserine from L-arginine. ³² The molecule has a long history as an inhibitor of PLP-dependent alanine racemases and has been used as an antibiotic for treatment of tuberculosis. The enzyme D-cycloserine A (DscA) is an arginine N-oxygenase (Figure 12.13). The N-OH-arginine scaffold is shown as the C=N-OH, oxime part structure. That oxime can equilibrate with a tautomer that has the C-NH-OH part structure with a single C-N bond. The next enzyme DcsB acts as a hydrolytic lyase, splitting the N-OH guanidine group out as N-hydroxyurea and the coproduct L-ornithine. Synthetic N-hydroxyurea, albeit genotoxic, has been used in sickle cell patients.

The N-hydroxyurea is next reacted with O-acetyl-L-serine via a PLP-dependent enzyme reaction cycle. The enzyme in the first part of its catalytic cycle acts as a classic β -eliminase, releasing acetate as it generates an enamino acid = PLP adduct. That is captured by the nucleophilic OH group of the N-hydroxyurea cosubstrate to generate an O-ureido-L-serine product.

The fourth enzyme in the pathway is a member of the ATP “grasp” family of kinases (see Chapter 15 for information on the ATP grasp enzyme family), creating a transient O-ureido-serine-phosphate (acyl phosphate) that is now set up for intramolecular ring formation as Pi departs with what had been the carboxylate oxygen of the O-ureido-L-serine. The product is a carbamylated form of D-cycloserine, with a novel N-O bond in an oxazolidinone ring. Action of a hydrolase on the carbamyl group yields ammonia, carbon dioxide and free D-cycloserine. This is a novel chemical strategy to create a ureido group and then generate a carbamyl-protected form of the product, hydrolysable on demand to the oxazolidinone heterocyclic warhead of the D-cycloserine antibiotic. The N-hydroxylation of arginine followed by hydrolase action to reveal the free N-hydroxyurea further illustrates the interplay between guanidine groups and ureas seen in the urea cycle enzymatic manipulations (Chapter 7). Analogously, the interplay between the ureido serine derivative, and intramolecular attack by the internal ureido nitrogen on the activated carboxyl group to give the cyclized N-carbamyl oxazolidinone shows the interrelationships between O-ureido groups and oxazolidinone heterocycles.

12.8 Nitric Oxide Synthase. Three-electron N-Oxidation of Arginine Guanidium Group

A second metabolic route from L-arginine through the arginine N-OH intermediate generates a very different fragmentation set of products, central to many aspects of human metabolism and physiology. ³³ Nitric oxide synthases (NOS) convert L-arginine and two molecules of O₂ in the presence of two molecules of the hydride donor NADPH to products citrulline, the nitric oxide radical and water (Figure 12.14). We shall turn to mechanistic considerations in Chapter 14 that deals with one- and two-electron oxidations of nitrogen atoms. Here, we observe that nitric oxide biology in humans was a landmark in unraveling signaling mechanisms.

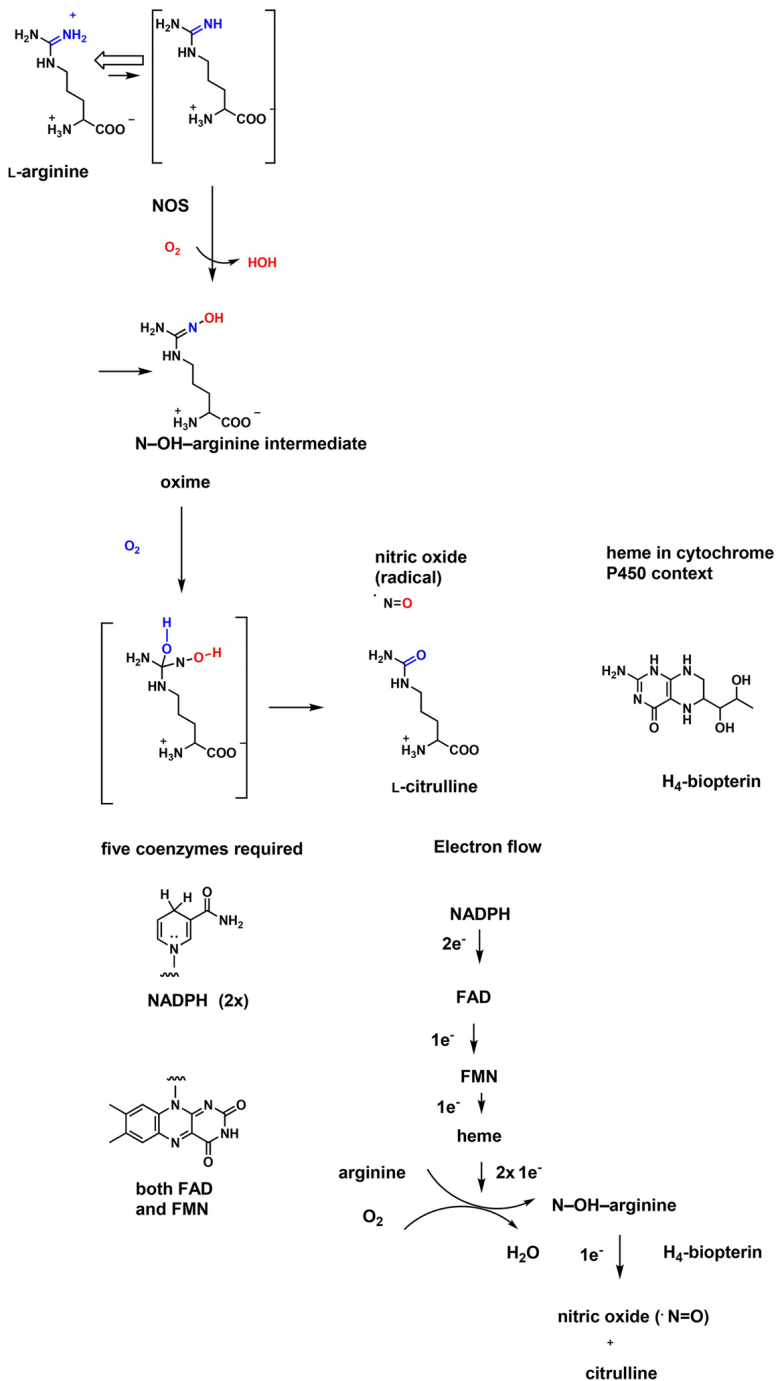


Figure 12.14 Conversion of arginine to nitric oxide by nitric oxide synthases: consumption of two O_2 via net three-electron oxidation of arginine and fragmentation to citrulline and nitric oxide radical. Five coenzymes are required, including two equivalents of NADH as external donors of two hydride equivalents

to bound FAD that acts as a $2e^- - 1e^-$ step-down electron conduit to FMN and then to the heme oxygenation catalyst. Tetrahydrobiopterin may be a one-electron transfer agent in the second oxygenation step.

In 1992 nitric oxide was designated the molecule of the year by Science Magazine. In 1998, three investigators shared the Nobel Prize in Physiology or Medicine for deciphering how it worked to lower blood pressure. Two lines of evidence, one from the discovery of an uncharacterized endothelial derived relaxation factor (lowering blood pressure) and the second from pathogens and nitrite metabolism, triangulated to allow identification of the nitric oxide radical, a gas at room temperature but with substantial aqueous and membrane solubility as the molecular entity.

Three different isoforms of nitric oxide synthase (NOS) are encoded in human genomes. ³³ One is the endothelial NOS (eNOS), responsible for NO production in lowering blood pressure. A second is neuronal NOS (nNOS) that generates NO to set in motion a number of neuronal programs in development and maintenance of excitable functions. Those two NOS are constitutively expressed. A third, inducible NOS (iNOS), is inducible in macrophages, where it sets off immunoprotective responses. The lifetime of the nitric oxide radical in different intracellular and extracellular microenvironments may be 1–10 seconds. We will note in Chapter 14 that this radical can react with a number of cellular constituents.

We have encountered nitric oxide previously in Chapters 3–5, e.g. as an intermediate in inorganic nitrogen redox pathways. In denitrification it is the second of three intermediates in that five-electron oxidation pathway. The traverse from NO_3^- to [NO] is a three-electron reduction, in the reverse direction compared with the three-electron traverse in the mammalian conversion of arginine to nitric oxide and citrulline.

12.8.1 Four Coenzymes Act *Ad Seriatim* to Funnel Electrons to a P450 Type Heme

All three human NOS isoforms employ *five coenzymes* to carry out the oxygenative fragmentation of arginine ⁴ (Figure 12.14). All five have nitrogen heterocycles. NADPH is the external electron donor: two equivalents serve up two hydride ions *ad seriatim* to tightly bound FAD coenzyme. The FAD paradigmatically serves as a two-electron–one-electron redox transformer, accepting electrons two at a time from NADPH. The FADH_2 form then transfers one electron at a time to a second bound flavin coenzyme, this one flavin mononucleotide (FMN). In turn the [FMNH] acts as “a one electron at a time” conduit to a bound heme cofactor (Figure 12.14). The calcium ion-activated protein calmodulin is an allosteric activator of the NOS isoforms, adding another level of regulation (calcium ion burst to NO bursts) to the NOS signaling regimes.

The heme microenvironment is that of a cytochrome P450 subtype (Chapter 11, Figure 11.13B). The Fe^{2+} in the equatorial plane of the heme tetrapyrrole macrocycle has a cysteine thiolate as bottom axial ligand (Figure 12.14). The top axial ligand will be occupied, successively, by two molecules of O_2 . Each O_2 will be reductively activated, split and transfer one of its two oxygens to the guanidine nitrogen in the first half reaction and then to form the incipient ureido carbonyl group of citrulline. Thus, the oxygenations occur at the heme coenzyme, not at the FAD coenzyme, in contrast to the *N*-hydroxylations of ornithine and lysine noted in siderophore biosynthesis earlier in this chapter.

The fifth coenzyme required for NO synthase catalytic cycles is the tetrahydro form of biopterin (BH_4) (Figure 12.14). BH_4 is a sparingly used form of the pterin ring system (Figure 12.15). It is an obligate cofactor, along with iron^{2+} in enzymatic hydroxylation of phenylalanine to tyrosine and then of tyrosine to

dihydroxyphenylalanine (DOPA; in neurotransmitter biosynthesis). The bicyclic pterin ring framework of BH₄ has functional analogies to the tricyclic isoalloxazine ring of flavins and can serve both as a two-electron–one-electron redox switch and react with O₂. In the three-electron catalytic cycle of nitric oxide synthases BH₄ is thought to play a role only in the transfer of the third electron.

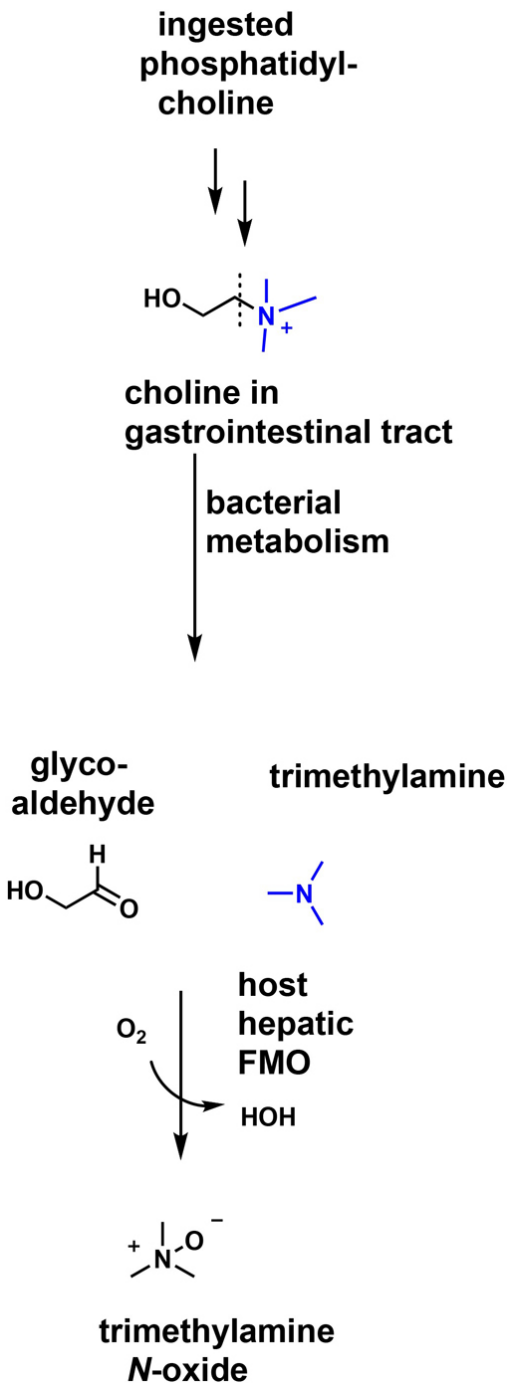


Figure 12.15 Trimethylamine oxide is produced by human liver FMO isozyme by *N*-oxygenation of the gastrointestinal tract bacterial metabolite trimethylamine. In turn, trimethylamine derives from dietary phosphatidylcholine *via* choline and a bacterial enzyme that splits choline to glycoaldehyde and trimethylamine in the

We take up the logic of high valent heme iron oxide species as mono-oxygen transfer agents to nitrogen, including weakly nucleophilic nitrogen, in Chapter 14. Similarly, we summarize the biology of the free radical metabolite nitric oxide in that chapter on one- vs. two-electron chemistry of nitrogen atoms in biology. The P450 type heme in nitric oxide synthases is mirrored by plant P450 *N*-oxygenases in plant nitrile biogenesis pathways described in the next chapter.

12.9 Human Flavin Monooxygenase (FMO) Isozymes

A second set of human *N*-oxygenases, in addition to the NOS isozymes above, comprise a small cluster of NADPH-utilizing FAD-dependent monooxygenases [flavin monooxygenases (FMOs)].³⁴ The family is conserved throughout the eukaryotic kingdom. There is a single isozyme in yeast but six in humans, all encoded on chromosome 1. In human liver the major isozyme is FMO3, while in intestine and kidneys it is FMO1.

The yeast enzyme is specific for thiol group oxidation and this may have been the ancestral function. In humans the five to six FMO enzymes have evolved to broad specificity oxygenases, acting on a variety of substrates with soft nucleophiles – nitrogen, sulfur, phosphorus – indicating a xenobiotic detoxification strategy.³⁵ These detoxifications are paralleled by a bevy of hepatic iron-containing cytochrome P450s that also carry out phase I drug metabolism (substrate hydroxylations) to increase the hydrophilicity of lipidic xenobiotic compounds so they can be excreted in the urine.

The mechanism of FMOs falls squarely in the NADPH–FAD oxygenase regime. Two-electron reduction of the bound FAD by substrate NADPH creates the FADH₂, oxygen-reactive form of the dihydroflavin. One-electron transfer to cosubstrate O₂ yields transiently the radical pair superoxide ion and flavin semiquinone [FADH] that collapse regiospecifically to the covalent FAD_{4a}–OOH. The conventional thought is that electron-rich sulfur, nitrogen or phosphorus atoms in cosubstrates attack the distal OH of the peroxyflavin, transferring it as an [OH]⁺ equivalent (although some have argued for one-electron pathways by homolysis of the weak O–O peroxy bond).

In addition to *N*-oxygenation of endogenous amine metabolites such as noradrenaline and adrenaline and tyramine, recent studies on the sequential bacterial and human metabolism of dietary choline, largely as ingested lipid phosphatidylcholine, has shown the involvement of hepatic FMO3. Bacteria in the gastrointestinal (GI) tract carry out a radical-based fragmentation of choline to glycoaldehyde and trimethylamine (Figure 12.15). Trimethylamine, as the free base, passes through the GI mucosa into the blood and *via* the portal vein reaches the liver. There it is an active substrate for the FMO3 isozyme and is quantitatively converted to trimethylamine-*N*-oxide. Studies by Hazen and colleagues⁵ have indicated that the product, trimethylamine oxide, modulates platelet responsiveness and thrombosis risk.

References

1. R. Winkler and C. Hertweck Biosynthesis of nitro compounds, *ChemBioChem*, 2007, **8**, 973-977.
2. J. H. Crosa and C. T. Walsh The genetics and assembly line enzymology of siderophore biosynthesis, *Microbiol. Mol. Biol. Rev.*, 2002, **66**, 223-249.
3. C. T. Walsh and T. Wencewicz Flavoenzymes: versatile catalysts in biosynthetic

- pathways, *Nat. Prod. Rep.*, 2013, **30**, 175-200.
4. C. Fang Mechanism of nitric oxide synthase regulation: electron transfer and interdomain interactions, *Coord. Chem. Rev.*, 2012, **256**, 393-411.
5. W. Zhu, J. Buffa, Z. Wang, M. Warrier, R. Schugar, D. M. Shih, N. Gupta, J. C. Gregory, E. Org, X. Fu, J. A. Didonato, A. J. Lusis, J. M. Brown and S. L. Hazen Flavin monooxygenase 3, the host hepatic enzyme in the metaorganismal trimethylamine *N*-oxide-generating pathway modulates platelet responsiveness and thrombosis risk, *J. Thromb. Haemostasis*, 2019, **16**, 1857-1872.
6. J. Olucha and A. Lamb Mechanistic and structural studies of the *N*-hydroxylating flavin monooxygenases, *Bioorg. Chem.*, 2011, **39**, 171-177.
7. J. L. Ronan, N. Kadi, S. A. McMahon, J. H. Naismith, L. M. Alkhalaf and G. L. Challis Desferrioxamine biosynthesis: diverse hydroxamate assembly by substrate-tolerant acyl transferase Des C, *Philos. Trans. R. Soc., B*, 2018, **373**10.1098/rstb.2017.0068.
8. J. Olucha, K. Meneely, A. S. Chilton and A. L. Lamb Two structures of an *N*-hydroxylating flavoprotein monooxygenase: ornithine hydroxylase from *Pseudomonas aeruginosa*, *J. Biol. Chem.*, 2011, **286**, 31789-31798.
9. V. Pohlmann and M. Marahiel Delta amino group hydroxylation of ornithine during coelichelin biosynthesis, *Org. Biomol. Chem.*, 2008, **6**, 1843-1848.
10. R. R. Jones and R. Bergman *p*-Benzyne. Generation as an intermediate in a thermal isomerization reaction and trapping evidence for the 1,4-benzenediyl structure, *J. Am. Chem. Soc.*, 1972, **94**, 660-661.
11. U. Galm, M. Hager, S. G. Van Lanen, J. Ju, J. S. Thorson and B. Shen Antitumor antibiotics: bleomycin, enediynes and mitomycin, *Chem. Rev.*, 2005, **105**, 739-758.
12. C. Zhang, E. Bitto, R. D. Goff, S. Singh, C. A. Bingman, B. R. Griffith, C. Albermann, G. N. Phillips and J. S. Thorson Biochemical and structural insights of the early steps in calecheamycin biosynthesis, *Chem. Biol.*, 2008, **15**, 842-853.
13. I. E. Sanderby, F. Geu-Flores and B. A. Halkier Biosynthesis of glucosinolates-gene discovery and beyond, *Trends Plant Sci.*, 2010, **15**, 283-290.
14. Y. C. Chen, E. Holmes, J. Rajniak, J.-G. Kim, S. Tang, C. R. Fischer, M. B. Mudgett and E. S. Sattely *N*-hydroxy-pipecolic acid is a mobile metabolite that induces systemic disease resistance in *Arabidopsis*, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E4920-E4929.
15. G. J. Gatto, T. B. Michael, N. L. Kelleher and C. T. Walsh Biosynthesis of pipecolate by RapL, a lysine cyclodeaminase encoded in the rapamycin gene cluster, *J. Am. Chem. Soc.*, 2006, **128**, 3838-3847.
16. J. Francke, I. Keishi and C. Hertweck Plasticity of the malleobactin pathway and impact on siderophore action in human pathogenic bacteria, *Chemistry*, 2015, **21**, 8010-8014.
17. J. Francke, I. Keishi, M. Ishida-Ito and C. Hertweck Nitro vs. hydroxamate in siderophores of pathogenic bacteria: effect of missing hydroxylamine protection in malleobactin biosynthesis, *Angew. Chem., Int. Ed.*, 2013, **52**, 8271-8274.
18. T. Szkudelski The mechanism of alloxan and streptozotocin action in B cells of the rat pancreas, *Physiol. Res.*, 2001, **50**, 527-546.
19. H. Lu, C. Emmanuel and H. Zhao CmlI is an *N*-oxygenase in the biosynthesis of chloramphenicol, *Tetrahedron*, 2012, **68**, 7651-7654.
20. A. J. Komor, J. J. Andrew, L. Que and J. D. Lipscomb Diiron monooxygenases in natural product biosynthesis, *Nat. Prod. Rep.*, 2018, **35**, 646-659.
21. K. Francis, S. Crystal, S. F. Nishino, J. C. Spain and G. Gadda The biochemistry of the metabolic poison propionate-3-nitronate and its conjugate acid 3-nitropropionate, *IUBMB Life*, 2013, **65**, 759-768.
22. A. J. Waldman, L. N. Tai, P. Wang and E. P. Balskus Heteroatom-heteroatom bond formation in natural product biosynthesis, *Chem. Rev.*, 2017, **117**, 5784-5863.

23. L. K. Wendy and A. T. Craig450 NocL in nocardicin A biosynthesis Role of the cytochrome, *J. Am. Chem. Soc.*, 2002, **124**, 8186-8187.
24. H. Ali, I. R. Marco, J. G. Nijland, P. P. Lankhorst, T. Hankemeier, R. A. L. Bovenberg, R. J. Vreken and A. J. M. DriessenA branched biosynthetic pathway is involved in the production of roquefortine and related compounds in *penicillium chrysogenum*, *PLoS One*, 2013, **8**(6), e65328.
25. B. LukeThe biosynthesis of glucosinolates: insights, inconsistencies, and unknowns, *Annu. Plant Rev.*, 2019, **2**(3), apr0708.
26. I. E. Sonderby, F. Geu-Flores and B. A. HalkierBiosynthesis of glucosinolates-gene discovery and beyond, *Trends Plant Sci.*, 2010, **15**, 283-290.
27. B. A. Halkier and G. JonathanBiology and biochemistry of glucosinolates, *Annu. Rev. Plant Biol.*, 2006, **57**, 303-333.
28. D. Galonic Fujimori, H. Sinisa, C. S. Neumann, M. Strieker, M. A. Marahiel and C. T. WalshCloning and characterization of the biosynthetic gene cluster for kutznerides, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16498-16503.
29. Y. L. Du, H. Hai-Yan, M. A. Higin and K. S. RyanA heme-dependent enzyme forms the nitrogen-nitrogen bond in piperazate, *Nat. Chem. Biol.*, 2017, **13**, 836-838.
30. K. D. Morgan, J. A. Raymond and K. S. RyanPiperazic acid-containing natural products: structures and biosynthesis, *Nat. Prod. Rep.*, 2019, **36**, 1628-1653.
31. G. Le Goff and O. JamalNatural hydrazine-containing compounds: biosynthesis, isolation, biological activities and synthesis, *Bioorg. Med. Chem.*, 2014, **22**, 6529-6544.
32. Y. Matoba, U. Narutoshi, M. Kudo and M. SugiyamaCyclization mechanism catalyzed by an ATP-grasp enzyme essential for d-cycloserine biosynthesis, *FEBS J.*, 2019, **287**, 2763-2778.
33. U. Forstermann and W. C. SessaNitric oxide synthases: *regulation and function*, *Eur. Heart J.*, 2012, **33**, 829-837.
34. S. Eswaramoorthy, B. B. Jeffrey, S. K. Burley and S. SwaminathanMechanism of action of a flavin-containing monooxygenase, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9832-9837.
35. M. Z. DanielAn overview of the mechanism, substrate specificities, and structure of FMOs, *Drug Metab. Rev.*, 2002, **34**, 503-511.

Nitriles, Cyanohydrins and Cyanogenesis

13.1 Hydrogen Cyanide and Cyanide Anion

Hydrogen cyanide (*aka* hydrocyanic acid) is a weak acid ($pK_a = 9.2$), it ionizes to about 1% at physiological pH values of 7.2 (Figure 13.1A). HCN melts at $-13\text{ }^{\circ}\text{C}$ to a colorless liquid and vaporizes to a colorless gas at $25\text{ }^{\circ}\text{C}$. The cyanide anion is a strong base and a strong carbon nucleophile. Historically, HCN has been known as a human poison for eons. The cyanide ion is a potent ligand for the cobalt ion in the aquo form of vitamin B₁₂, generating cyanocobalamin. The cyanide ion is also a tight binding axial ligand to Fe^{II} in heme groups in proteins such as catalase.

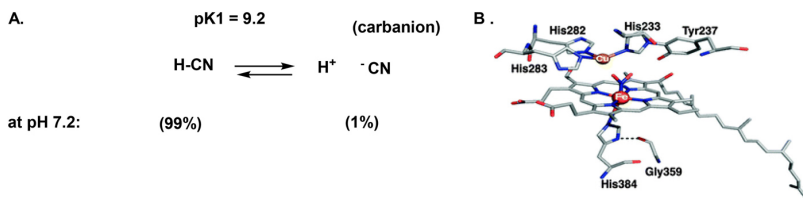


Figure 13.1 (A) HCN dissociates to a cyanide ion and a proton with a pK_a of 9.2; (B) The prime target for cyanide toxicity is inhibition of the terminal enzyme in mitochondrial respiratory chains, cytochrome C oxidase which contains both heme-a₃ and a histidine-coordinated copper atom. Reproduced from [ref. 25](#) with permission from the Royal Society of Chemistry, Copyright 2015.

The most susceptible and consequential toxic event appears to be axial ligation to the iron in the heme group (cytochrome a₃) in mitochondrial cytochrome C oxidase (Figure 13.1B), the terminal enzyme in respiratory electron transport chains, funneling four electrons into O₂ as it undergoes reductive fragmentation to two H₂O product molecules. 1 Blockade of this enzyme turns off O₂ reduction, energy production and ATP synthesis and results in metabolically active tissues, such as muscles and nerves, running out of energy and dying. At lethal doses of cyanide (*e.g.* K⁺ CN⁻), approximately 100–200 mg for a 70 kg human adult, death can be rapid.

Sublethal doses can be detoxified in humans by capture by electron deficient sulfur ions released by action of mercaptopyruvate lyase. 2 As shown in Figure 13.2, enzyme-catalyzed C₃–S bond cleavage generates the enolate of pyruvate and a transient enzyme cysteine–S–SH persulfide linkage. That terminal persulfide sulfur atom is

electrophilic and capturable by a cyanide ion acting as a carbanion. The result is the thiocyanate ion (SCN^-), with the negative charge on sulfur. This is a stable, nontoxic derivative of the cyanide ion.

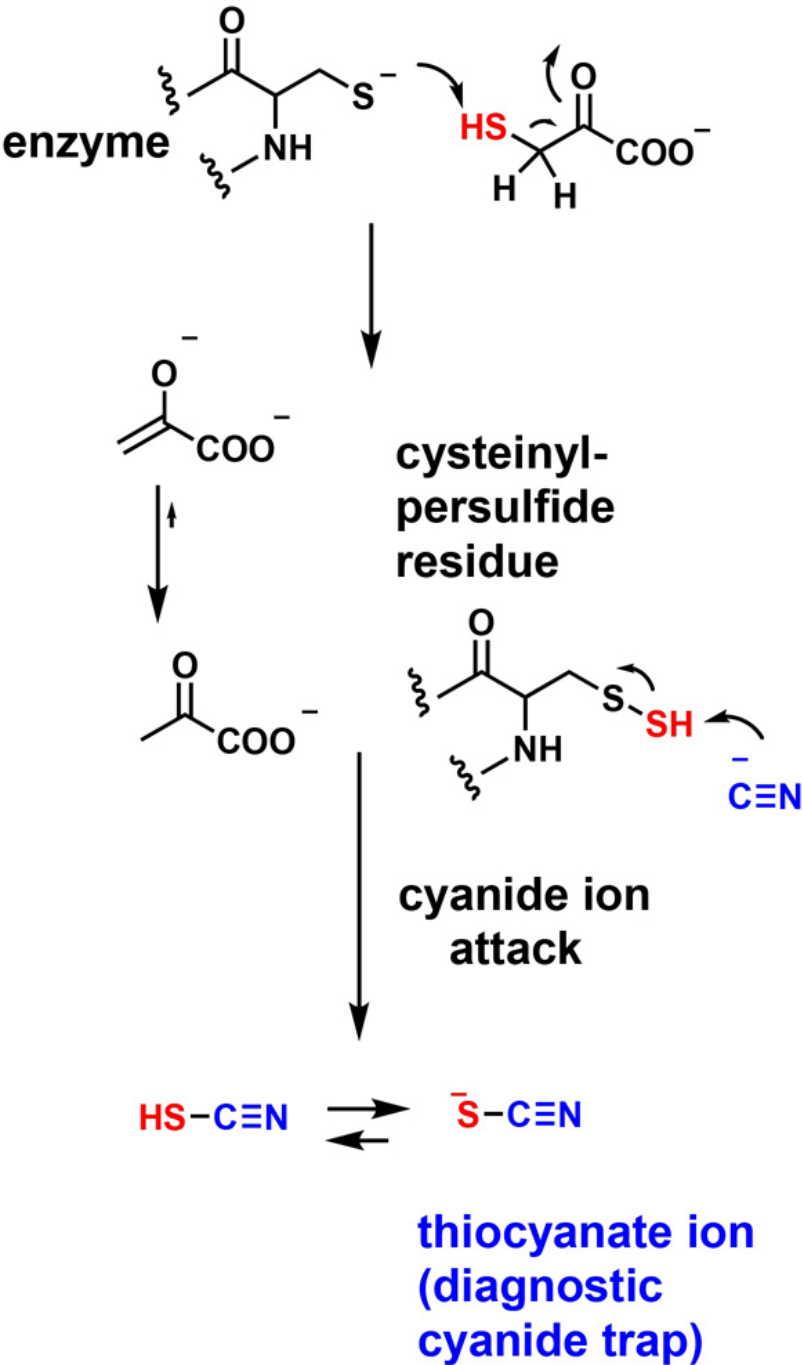


Figure 13.2 A diagnostic test for cyanide ion: capture of an enzyme–cysteine–SSH terminal sulfur by cyanide ion generates a thiocyanate anion. Mercaptopyruvate lyase provides a catalytic example.

In a sense the cyanide ion is the simplest organic nitrogen structure. The nitrogen is triply bonded to the lone carbon. In fact, the cyanide ion is isoelectronic to N_2 , as well as to CO and the nitrosonium ion (Figure 13.3). Despite its toxicity as a strong axial ligand to ferrous ion in heme proteins, cyanide is a natural metabolite, typically generated by plants as a defense agent. Also, certain bacterial hydrogenases that contain both nickel and nonheme iron atoms at their catalytic centers use cyanide as a proximal ligand for iron: we will note the conversion of the common amino acid tyrosine to a cyanide ion for subsequent iron ligation.

isoelectronic compounds



cyanide ion



dinitrogen

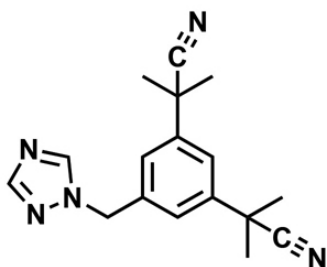
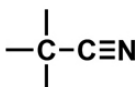


carbon monoxide



nitrosonium ion

organic cyano compounds = nitriles



**astrazole
(breast tumor drug)**

Figure 13.3 Compounds isoelectronic with cyanide ions; organic nitriles with the breast tumor drug astrazole as an example.

Plants typically store the cyanide equivalent in an inactive precursor form, as

cyanogenic glycosides, some two enzymatic steps away from free cyanide ion. 3·4 It appears that some 3000 plants have the genes needed to make cyanogenic glycosides. The cyanogenic portions are typically cyanohydrins, addition products of cyanide ion to aldehydes. These are metabolically and chemically distinct from organic compounds with stable C–CN bonds, known predominantly as nitriles (Figure 13.3). Unlike cyanohydrins (carbonyl adducts), organic nitriles are not ready progenitors of the cyanide ion and are not toxic *per se*. In this chapter, we explore nitrile biosynthesis, storage as cyanohydrin precursors and multiple routes to cyanogenesis.

13.2 Plant Cytochrome P450s Convert Amino Acids to Cyanohydrin Glycosides

13.2.1 The Subclass of Cyp79 Heme *N*-Oxygenases Convert Amino Acids to Aldoximes

A variety of plants have evolved a subcategory of heme-containing cytochrome P450 monooxygenases, designated Cyp79 subfamily members, to display *N*-oxygenation capability towards amino acids. The specificity towards polar water-soluble amino acids is atypical for plant P450s that normally bind and oxygenate hydrophobic substrate scaffolds.

The Cyp79 family carry out two consecutive *N*-hydroxylations on a substrate, such as tyrosine, 5 shown in Figure 13.4A. These follow the general pattern for *N*-oxidation discussed in Chapter 12 except that it is the α NH₂ group not a side chain amine that gets *N*-oxygenated [in a typical reduced nicotinamide-adenine dinucleotide phosphate (NADPH) to flavin adenine dinucleotide (FAD) to heme electron transfer sequence with FAD acting as the $2e^- - 1e^-$ step-down redox switch to heme in the enzymes]. Two such monooxygenative catalytic cycles yield the *N,N*-dihydroxy tyrosine that loses water (presumably nonenzymatically) to the *C*-nitroso product. The *C*-nitroso group is then isomerized to the oxime as decarboxylation provides the driving force for the nitroso to oxime isomerization. Purified CYP79A1 from sorghum catalyzes both of these two-electron oxidative steps with two attendant water molecule products, plus the subsequent decarboxylative isomerization of the *C*-nitroso group to aldoxime (Figure 13.4).

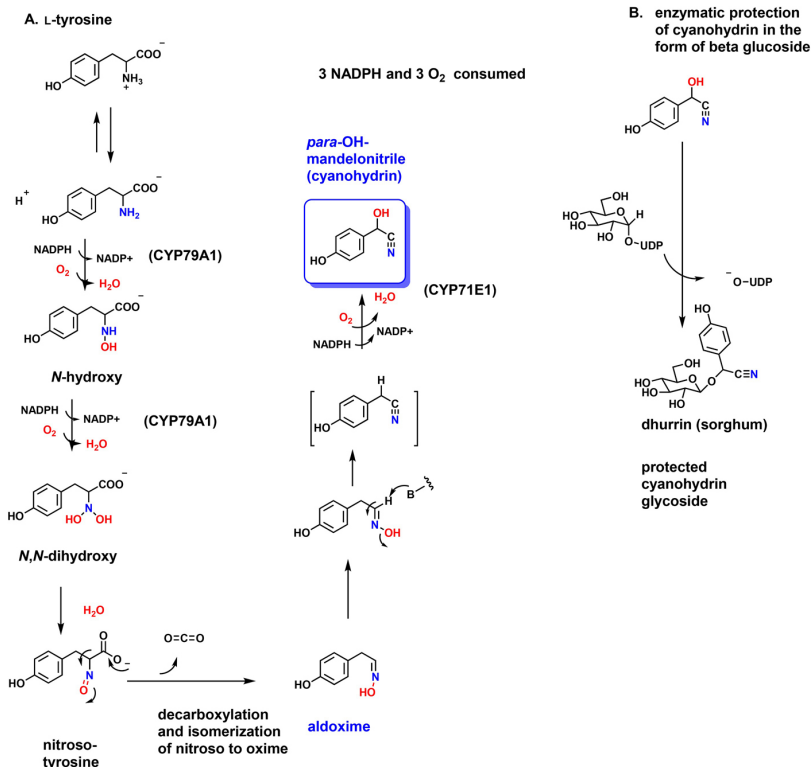


Figure 13.4 (A) Tandem action of CYP79A1 and CYP71E1 convert tyrosine to *para*-hydroxy-mandelonitrile. (B) Cyanohydrins are protected as glucosides *in planta*.

13.2.2 From Aldoxime to Trapped Cyanohydrin Acetals

Next, a second plant cytochrome oxygenase, **5** from the CYP71E1 subfamily, catalyzes a third oxygenation. The oxime is desaturated to the triple bond in the newly formed nitrile group as the C₂ carbon is hydroxylated. Three successive P450 oxygenase cycles have carried out a six-electron oxidation of the amino acid to the cyanohydrin scaffold of *para*-hydroxymandelonitrile (Figure 13.4A). The cyanohydrin, not an unmodified organic nitrile, is the direct product of this six-electron redox process.

The cyanohydrin can decompose nonenzymatically to the aldehyde and free cyanide. However, it is typically captured *in planta*, before that slow decomposition, by a glycosyl transferase. In the case depicted in Figure 13.4B the glycosylation substrate is the classic glycosyl donor UDP glucose. An early transition state favors a glycosyl C₁-oxocarbenium ion as the electrophile, capturable by the oxygen of the cyanohydrin. The end product is the *O*-glucosylated cyanohydrin, a chemically stable acetal linkage. That glycosylated cyanohydrin from sorghum is known as dhuririn.

With its glycosidic acetal linkage dhuririn and related cyanogenic glycosides are metabolically stable precursors, not decomposing without enzymatic assistance to the free cyanide ion. They can be prestored in plant tissues awaiting an alarm signal (as phytoanticipins) that a predator is near, for example nibbling on and damaging leaves of the plant. **6**

Additional amino acids that can be oxidatively converted to the one-carbon

shortened cyanohydrins in other plants include phenylalanine, valine and isoleucine to give such cyanogenic glycosides as amygdalin, linamarin and lotaustralin 7·8 (Figure 13.5). Prunasin and sambunigrin have the same glucose protectant moiety but different chirality at the cyanohydrin. Amygdalin is a prunasin-derived diglycoside. Tetraphyllin B has an unusual nonproteinogenic 2-(4-OH)-cyclopentenylglycine amino acid moiety as a cyanide precursor.

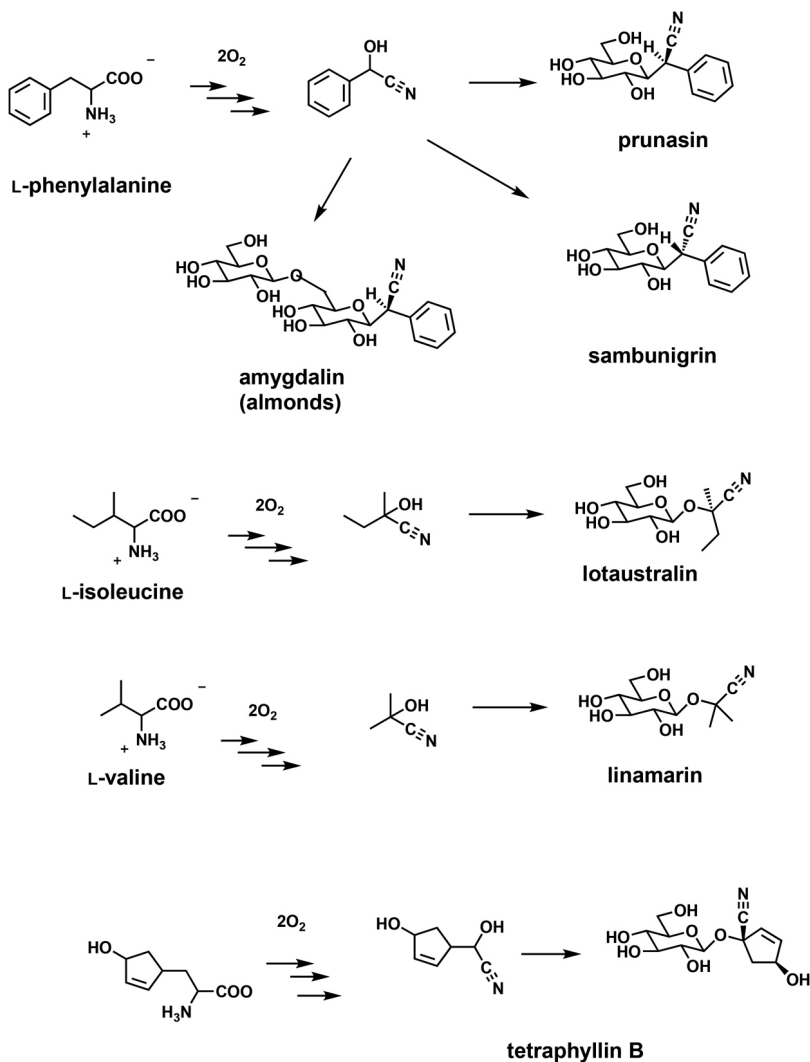


Figure 13.5 Additional plant cyanogenic glycosides.

13.3 Two-step Cyanogenesis from the Cyanogenic Glycosides

Among the thousands of plants that store cyanogenic glycosides, the controlled

defensive release of free cyanide ion occurs in a two-step enzyme-catalyzed process. The ultimate release of cyanide ion is initiated by dissolution of an internal membrane compartment, allowing *beta*-glycosidases to gain access to the stored cyanogenic glucosides. **Figure 13.6** shows the acetal linkage of prunasin undergoing catalyzed hydrolysis to the cyanohydrin *para*-hydroxyphenyl-mandelonitrile. **8** This uncovered cyanohydrin can undergo leisurely nonenzymatic decomposition to cyanide and *para*-OH-benzaldehyde but that rate is too slow to be useful to the plant in an exigency. Instead, plants have α -hydroxynitrile lyases that accelerate the decomposition into the toxic cyanide ion and the aryl aldehyde. The hydroxynitrile lyases are members of the $\alpha\beta$ -hydrolase superfamily with a classic aspartate, histidine, serine (Asp, His, Ser) catalytic triad. **9** Structural studies have indicated that the active site side chain of a serine residue abstracts the O–H proton to initiate conversion to the aldehyde with expulsion of a cyanide ion (**Figure 13.6**).

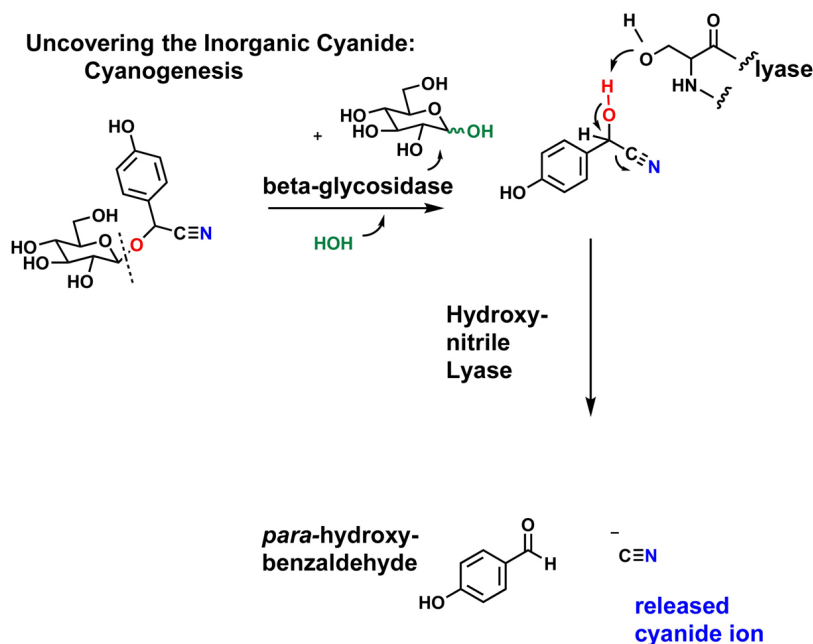


Figure 13.6 Activation of the cyanide ion from the cyanohydrin glycosides requires two enzymatic steps. As shown for prunasin, a glycosidase and then a hydroxynitrilase lead to a free cyanide ion and *para*-hydroxybenzaldehyde.

13.4 Recapture of Cyanide by Cyanoalanine Synthase and Conversion to Asparagine or Aspartate

Once a plant has catalyzed hydrolysis of cyanogenic glucosides and unraveling of the consequent cyanohydrin, it may be subject to a transient high local concentration of HCN or CN^- that could prove toxic to the plant hemoproteins (e.g. cytochrome C oxidase). To cope with that liability plants elaborate a pyridoxal phosphate (PLP)-dependent *beta* cyanoalanine synthase **10** to act as an internal cyanide detoxification

catalyst (Figure 13.7).

**Recapturing the Cyanide Ion:
Detoxification and Return to
The Amino Acid Pool**

Beta Cyanoalanine Synthase

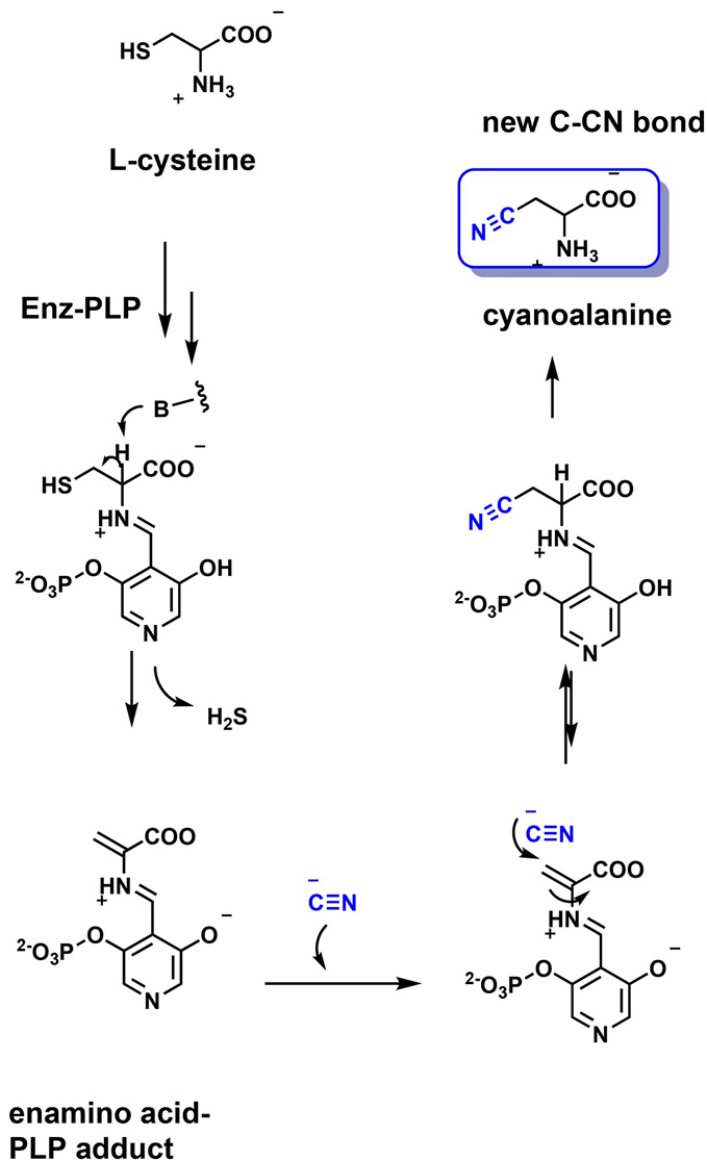


Figure 13.7 A cyanide ion is recaptured and returned to the metabolic pool as

cyanoalanine by a pyridoxal phosphate-enzyme that catalyzes a net beta replacement of the -SH in cysteine by the cyano group in the cyanoalanine product.

The cosubstrate is the central amino acid L-cysteine. On binding to this synthase, cysteine undergoes the normal transaldimination to yield the Cys-PLP aldimine adduct. Now, the C₂-hydrogen on the cysteine scaffold is kinetically acidic because the resultant carbanion is delocalized through the PLP framework. PLP thus acts as usual as a carbanion-stabilizing electron sink. That C₂ carbanion can be used as a neighboring pair of electrons to expel H₂S, cleaving the C₃-S bond as the three-carbon enamino acid-PLP adduct is formed. This adduct is electrophilic at C₃ by virtue of the conjugation into the PLP electron sink.

The cosubstrate HCN has about one part in a hundred as the cyanide ion at pH 7.2 (pK_a for HCN = 9.2) and that anion acts as a carbon nucleophile on C₃ of the enamino-PLP adduct as shown in [Figure 13.7](#). This is the C₃-CN bond forming step. All the HCN can be drawn off by the rapid deprotonation equilibrium with the cyanide anion. Reprotonation at C₂ of the cyanoalanyl-PLP carbanion adduct is the final step before release of free cyanoalanine. Cyanoalanine is nontoxic because the cyano moiety is covalently attached to the amino acid framework. It is a stable organic nitrile, not a labile cyanohydrin.

To bring the cyanide carbon back into the central cellular amino acid pool there are two variants of nitrile hydratases. [4·9·11](#) One class of nitrile hydratases catalyzes the addition of water into the cyano group of cyanoalanine. The initial isoamide tautomer isomerizes to the familiar amide group of asparagine ([Figure 13.8](#)). Alternatively, some nitrilases carry out two consecutive hydrations, generating bound asparagine as an intermediate but then hydrating its amide one more time to ammonia and aspartate ([Figure 13.8](#)). Either way, the nitrile carbon is fixed either as carboxamide or carboxylate as C₄ of Asn or Asp.

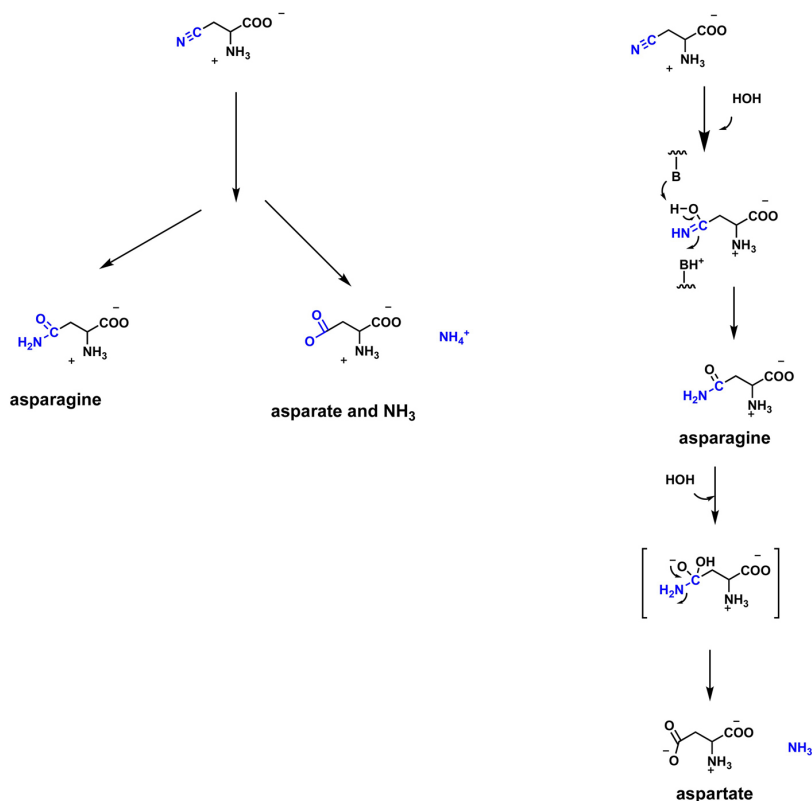


Figure 13.8 Cyanoalanine nitrile hydratases may add one equivalent of water to the nitrile group to generate the amide in asparagine or may go on to a second water addition to the amide to release free NH_3 + and aspartate.

13.5 Toyocamycin: Nitrile Formation and Reduction

Additional organic nitriles that are not cyanohydrins are known microbial metabolites. Some strains of bacteria make pyrrolopyrimidines, also known as 7-deazapurines. [12](#) They proceed by way of 7-cyano-7-deazapurine intermediates. Starting with the identification of toyocamycin ([Figure 13.9](#)) in 1956, more than 30 members of this heterobicyclic natural product family have been characterized. Also, among the modified nucleobases in certain tRNAs is the 7-deazaguanine framework of queosine ([Figure 13.9](#)). There are many unusual aspects to the 7-deazaguanine assembly strategies including the ring contraction–extrusion reaction by a radical-SAM enzyme QueE. [13](#) Readers are directed to publications from the Bandarian laboratory over the past decade to get insights into the full pathway logic. [12](#)[13](#)

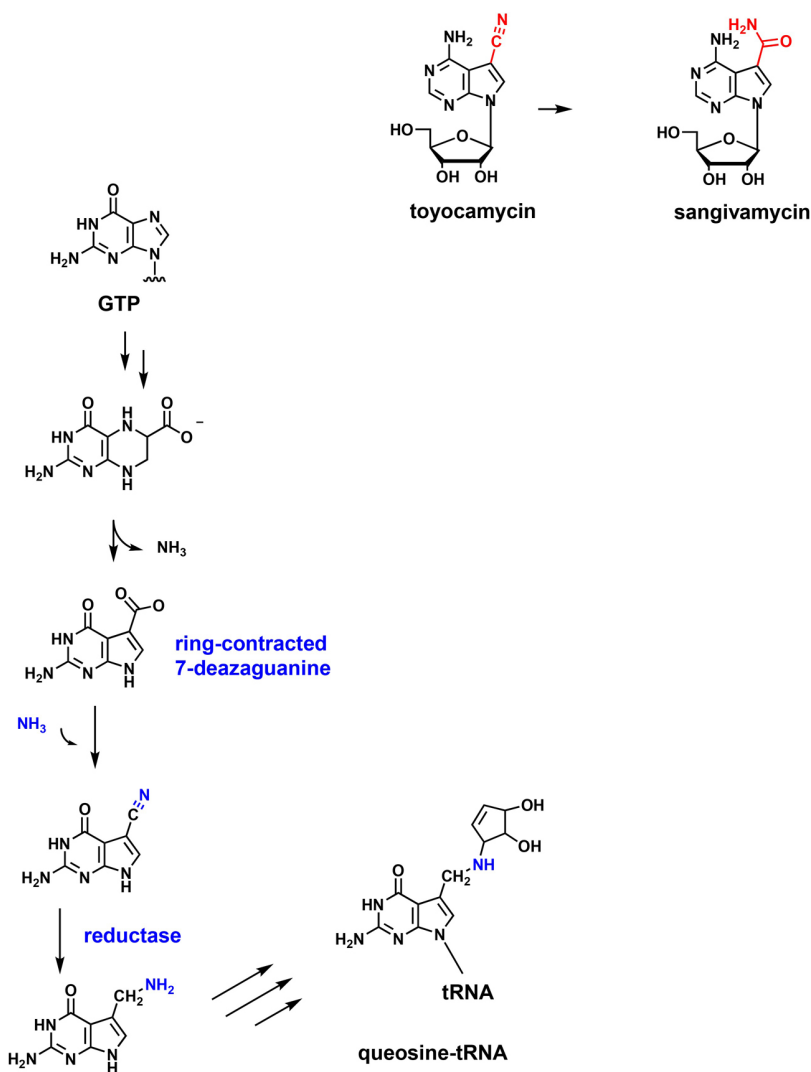


Figure 13.9 Conversion of GTP to 7-deazaguanines and manipulation of the cyano group to amine. Note that during ring contraction, free ammonia is generated as the extruded carbon remains at the level of a carboxylate. In the next step ammonia addition leads to the exocyclic nitrile group.

Here we focus only on the formation of the nitrile in prequeosine and its reduction to the corresponding amine. Conversion of the carboxylate to an amide and then on to the nitrile is mediated by iterative ATP-dependent logic. ¹⁴ Two ATPs are spent to carry out the sequential activation of the carboxylate to the amide and the amide to the nitrile.

Activation of the carboxylate as an acyl-AMP followed by amination from ammonia as cosubstrate (Figure 13.10) generates the amide, much as in asparagine synthetase. A second molecule of ATP is likewise cleaved, presumably by *O*-phosphorylation of the isoamide and then collapse of the tetrahedral adduct with expulsion of inorganic phosphate as the nitrile forms.

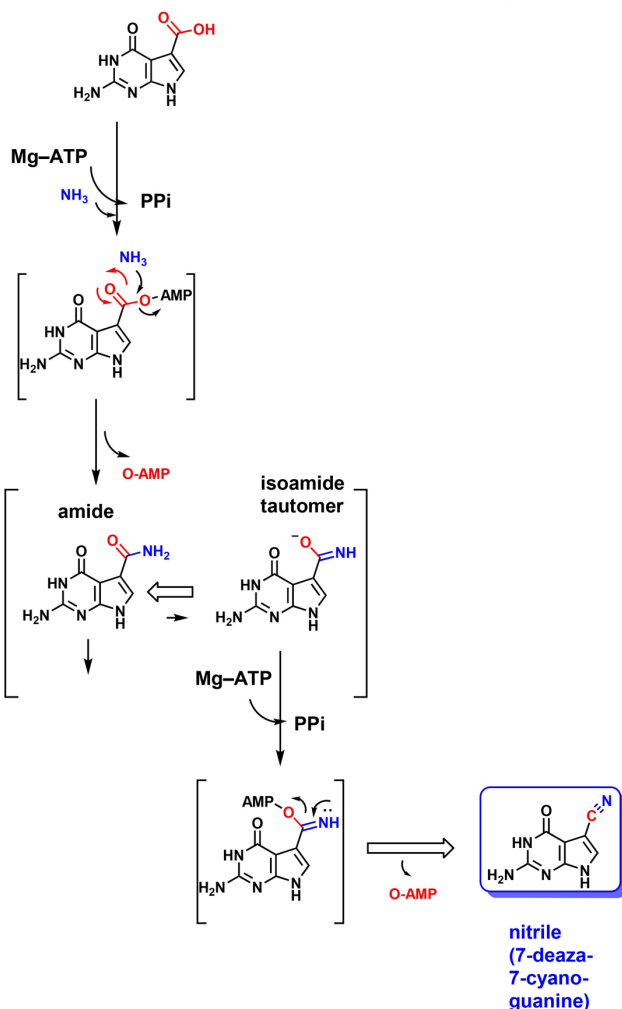


Figure 13.10 Conversion of 7-deazaguanine acid to amide to nitrile. Each of the two steps requires cleavage of an ATP to AMP and PPI to activate the oxygen atom to depart as part of an AMP-O leaving group.

This second step is akin to activation of citrulline for nucleophilic attack. In arginine biosynthesis. The difference here is that the proposed mechanism involves the lone pair of the imine nitrogen of the O-AMP adduct acting as neighboring group to expel AMP along with what had been the carboxamide nitrogen – a difficult C-O cleavage step in the absence of ATP activation (Figure 13.10).

The next biosynthetic step is enzymatic reduction of the nitrile to an amine, catalyzed by QueF. This is a net four-electron reduction (Figure 13.11). When coupled with the previous enzyme the net two-step conversion is of a carboxylate group to a primary amine by way of the intermediate nitrile. QueF from *Bacillus subtilis* has been crystallized and its structure determined by X-ray analysis, ¹⁵ giving insight into the reaction mechanism of the double reduction. As depicted in Figure 13.11, the reductase has an active site cysteine residue that is proposed to add into the nitrile

(acting as electrophile at carbon) to give a covalent enzyme thioimide intermediate. That is the framework proposed to accept a hydride ion from bound NADPH. The spent NADP^+ departs while the resultant thiohemiaminal covalent adduct could dissociate as shown to the free cysteine thiolate side chain and the aldimine form of the 7-deazaguanine. That aldimine is an electrophilic carbon capable of undergoing another hydride transfer to the imine carbon by the second NADPH as shown. The released product, the 7-amino-7-deazaguanine, has undergone two consecutive hydride transfers for a net four-electron reduction. That amine is further metabolized on the way to the queosine tRNA maturation.

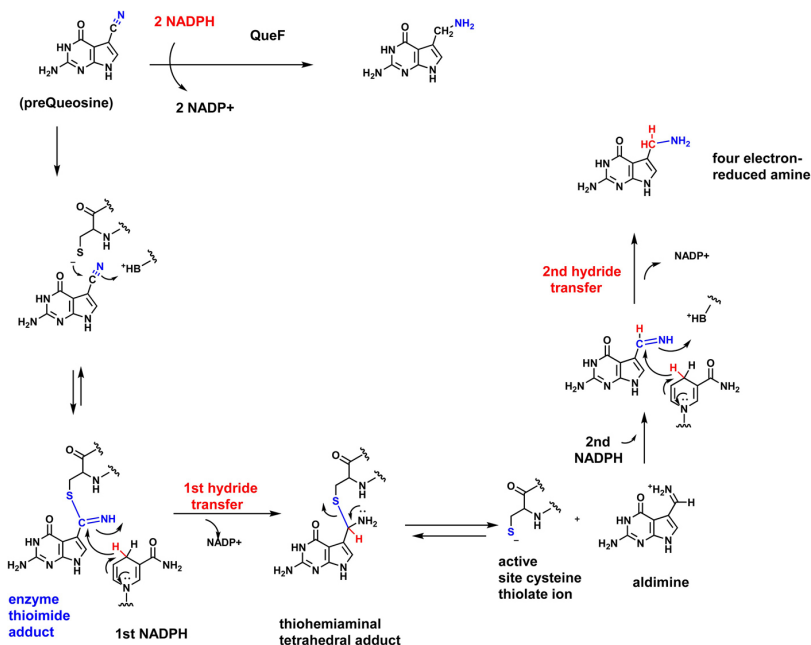


Figure 13.11 Conversion of 7-deazaguanine-nitrile to 7-deazaguanine amine: a net four-electron reduction mediated by two consecutive hydride transfers from two NADPH molecules. The proximal species reduced are: first a covalent enzyme thioimide adduct, second a free deazaguanine aldimine.

The two-enzyme conversion of 7-deaza-7-carboxy guanine to 7-deaza-7-amino guanine is a costly conversion in terms of cellular energy. Two ATPs are spent to take the carboxylate anion to the nitrile. Then, two equivalents of the reducing currency NADPH are spent for the four-electron reduction of nitrile to primary amine. The process is revealing of the logic of acyl group nitrogen chemistry and the energetic costs of those metabolic steps.

13.6 Additional Cyanogenesis Strategies

There is one additional plant-based cyanogenesis route and two known bacterial routes to generate free cyanide that reveal distinct strategies. The other plant-based pathway to cyanide is a byproduct of generation of the fruit-ripening hormone ethylene. The two bacterial routes take glycine or tyrosine respectively and convert the CH-NH_2

linkages to the triply bonded $C\equiv N$ nitrile group as cyanide is expelled during scaffold fragmentation. All three are oxidative fragmentative transformations of substrate amino acid scaffolds and all three probably involve nitrogen radicals at some point during the respective enzyme catalytic cycles, bringing organic nitrogen one-electron chemistry into focus.

13.6.1 Enzymatic Oxidative Fragmentation of 1-Amino-1-carboxy-cyclopropane to Ethylene, Cyanide Ion and Carbon Dioxide

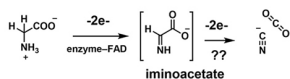
Ethylene has been known for decades as an endogenous fruit ripening hormone in plants. This simple C_2H_2 olefin diffuses to receptor isoforms, thought to contain redox active metals where the ethylene can form metal-olefin complexes to trigger downstream metabolic and genetic responses in different subcellular compartments in target cells.

We limit our focus to the specific reaction that releases ethylene and cyanide simultaneously. The substrate is an unusual cyclopropane containing amino acid known as 1-aminocyclopropane-1-carboxylate (ACC) [16·17](#) ([Figure 13.12A](#)). It originates from *S*-adenosylmethionine (SAM) by action of an unusual PLP-dependent enzyme. The SAM=PLP aldimine adduct as usual can form a stabilized C_2 -carbanion. In this adduct the C_2 carbanion can act as *intramolecular* carbon nucleophile on C_4 , electrophilic due to its adjacency to the cationic sulfonium group in SAM. This results in C_4-S^+ bond cleavage as a new C_2-C_4 bond forms. The released products are thiomethyladenosine and the ACC cyclopropane amino acid ([Figure 13.12A](#)).

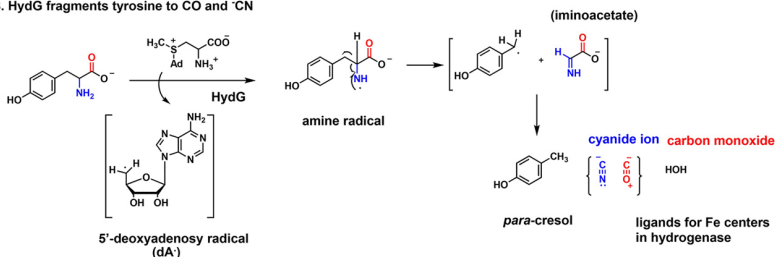
A. aminocyclopropane carboxylate oxidase



2. HCN synthase (glycine-fragmenting)



3. HydG fragments tyrosine to CO and CN^-



B. ACC scaffold falls apart to cyanide ethylene and carbon dioxide

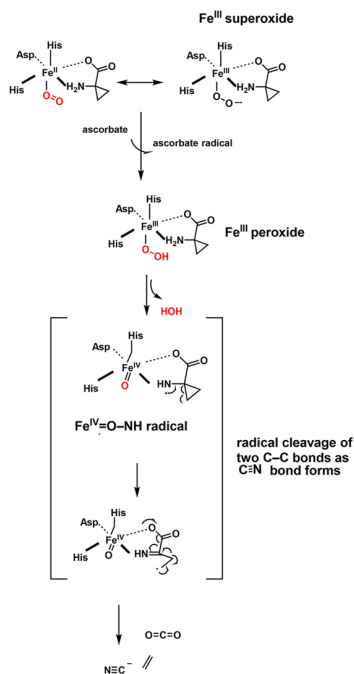


Figure 13.12 Three additional radical fragmentations that produce cyanide ion: (A) amino-cyclopropyl carboxylate (ACC) oxidase, HCN synthase, HydG fragmentation of tyrosine to CO and CN^- as ligand for the iron in hydrogenase. (B) Mechanistic proposal for O_2 -induced fragmentation of the ACC scaffold to cyanide ion, CO_2 and ethylene.

At first glance ACC does not appear to be an obvious carrier of latent ethylene,

which derives biologically from C₃ and C₄ of ACC, and thereby from C₃ and C₄ of methionine. However, the nonheme iron enzyme ACC oxidase uses O₂ to induce a net fragmentation of ACC 18 (Figure 13.12B). Two C–C bonds collapse as the single C₂–NH₂ bond in amine linkage is converted to triply bonded cyanide ion. Metaphorically, the ACC scaffold is blown up in a controlled room temperature oxidation. Neither of the two atoms in O₂ end up in ethylene, cyanide or carbon dioxide products. Instead, two molecules of water are generated, reflecting a net four-electron reduction of O₂ as ACC undergoes a fragmenting four-electron oxidation (Figure 13.12B). However, the O₂ is proposed to be utilized to create a high valent Fe^{IV}=O oxidant. 18 One-electron transfer from the ligated ACC substrate to iron is proposed to initiate the fragmentation. Ethylene is the desired hormone product. Cyanide seems to be a byproduct of the oxidative fragmentation mechanism. It can be detoxified by the cyanoalanine synthase route noted in the previous section.

13.6.2 Two Bacterial Amino Acid Fragmentations to Cyanide

Reactions 2 and 3 of Figure 13.12A show the stoichiometry for two bacterial enzymes that respectively fragment glycine or tyrosine to cyanide ions. The glycine cleavage enzyme has been termed HCN synthase. It is a membrane-bound enzyme in such bacteria as *Pseudomonas aeruginosa* and has not been dissected mechanistically other than a report that it requires FAD. 19 The redox nature of the fragmentation certainly indicates that redox cofactors, perhaps both organic and inorganic, would be required.

13.6.2.1 Glycine-fragmenting HCN Synthase

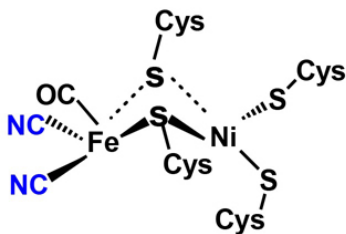
Classical amino acid dehydrogenases and oxidases have in common a net two-electron substrate oxidative half reaction. For glycine that would generate nascent iminoacetate and a reduced flavin coenzyme [FADH₂ or reduced flavin mononucleotide (FMN_{H2})]. The second half reaction, reoxidation of the dihydroflavin, would then be different. For dehydrogenases a physiological electron acceptor other than O₂ would be kinetically competent (e.g. cytochromes). For flavoenzyme oxidases, O₂ would be reduced by two electrons to H₂O₂. The nascent iminoacetate could be released and rapidly undergo attack by water on the imine carbon with forward decomposition to glyoxalate and NH₃.

The pseudomonal HCN synthase on the other hand, assuming it follows the prototypic first half reaction, must hold onto the nascent iminoacetate and oxidize it by a further two electrons, perhaps by single-electron transfer with radical intermediates. In so doing, the C₁–C₂ bond would fragment as CO₂ forms. Removal of the third and fourth electrons from the starting glycine scaffold turns the C₂=NH₂ + double bond into the cyanide ion triple bond. Unlike the bacteria discussed below in the apparently similar hydrogenase maturase G (HydG) fragmentation, pseudomonads do not contain hydrogenases and do not use the cyanide as an active site ligand and so the physiological purpose of the HCN synthase remains obscure. (The four electrons removed from glycine could be passed along membrane electron transport chains down to O₂ and thus serve as an energy source. The toxic cyanide product would need to be detoxified.)

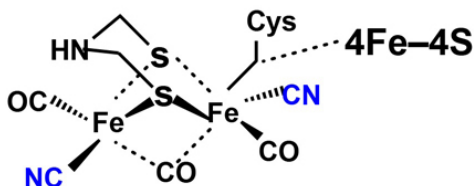
13.6.2.2 Enzymatic Fragmentation of Tyrosine to Provide CN and CO Ligands for Hydrogenases

Bacterial hydrogenases have been known for almost 100 years, largely contained in

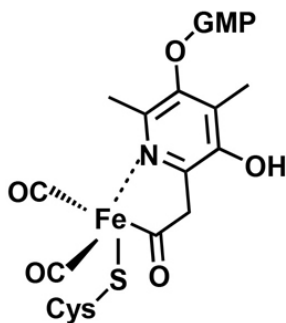
anaerobic bacteria since the Fe-S clusters in the hydrogenase active sites are typically oxygen labile and are rapidly inactivated in air. 20·21 An enormous amount of primary and review literature has accreted over the decades. Classifications of various hydrogenases are based on; (1) functional directionality of either H₂ uptake or H₂ evolution among producers in mixed bacterial communities, (2) the nature of the redox active iron clusters or (3) the presence of nickel-iron hybrid metal clusters. The ligand sets around the redox active irons have also been determined to high resolution by structural studies. Figure 13.13 shows the ligand sets for each of the Fe-only hydrogenases, for the Fe-Fe subgroup and for the third, Ni-Fe, subgroup. 20



[Ni-Fe] hydrogenase



[Fe-Fe] hydrogenase



[Fe] hydrogenase

Figure 13.13 Active site ligands for the three subclasses of hydrogenases: Fe-Ni hydrogenase, Fe-Fe hydrogenase and Fe hydrogenase.

All three of the active-site irons have carbon monoxide ligands to iron: two in the Fe

hydrogenase subgroup, one at each iron in the Fe-Fe variant with a bridging CO ligand. In the Ni-Fe variant there is one carbon monoxide and two cyanide ligands. Cyanide is absent from the Fe hydrogenase but present in one copy in the Fe-Fe hydrogenases. We do not focus here on the catalytic mechanisms of these three hydrogenase classes but limit our discussion to the known source of CO and CN in the Ni-Fe and Fe-Fe hydrogenases.

The joint provision of both carbon monoxide (CO) and cyanide ($-\text{CN}$) from a single tyrosine substrate molecule is the function of the HydG enzyme (Figure 13.12A). HydG also requires a molecule of *S*-adenosylmethionine. In contrast to its role in ACC formation, in HydG catalysis SAM undergoes one electron input and homolysis of the CH_2-S^+ bond to create a 5'-deoxyadenosyl fragment radical [dA]. The [dA] radical is a sufficiently robust one-electron oxidant that it can abstract hydrogen atoms [H] from nearby cosubstrates bound in the active site. 22:23

In this case abstraction of an [H] from bound cosubstrate tyrosine's amino group yields the amine radical noted in Figure 13.12A. That radical can rearrange as shown to cleave the C_2-C_3 bond of tyrosine, transiently creating the *para*-cresol radical and an iminoacetate fragment. Note that iminoacetate is the proposed intermediate above in the bacterial HCN synthase catalytic cycle, so the likely paths converge for these two cyanide-producing catalysts. How the iminoacetate is then cleaved to cyanide and CO is not understood at present. Note that the cyanide ion and carbon monoxide are isoelectronic. Both have a carbon attached to a heteroatom (O or N) by a triple bond (Figure 13.12). How HydG makes CO and H_2O during release of cyanide while HCN synthase makes CO_2 and a cyanide ion are likewise not apparent.

The result of HydG action is to provide equal amounts of carbon monoxide and cyanide in the local vicinity of maturing hydrogenase iron-iron and iron-nickel clusters. These two isoelectronic ligands are key tuning ligands for action of these two major subclasses of hydrogenases.

13.7 Biological Formation of Isonitriles

Isonitriles, isomeric with nitrile groups, are known metabolites but exceedingly rare in microbial metabolism. As denoted in Figure 13.14A, whereas nitriles have a terminal triply bonded nitrogen and a carbon with an R group substituent on the carbon atom, the isonitriles have reverse connectivity. The carbon is the terminal atom, the nitrogen is bound to the R group. In isonitriles then, the terminal carbon bears a formal negative charge while the nitrogen is quaternized and bears a positive charge. A neutral carbenoid resonance contributor to the isonitrile structure can be written. Indeed, isonitriles react abiotically as carbenes, with nucleophiles, electrophiles and radicals. They can be labile in aqueous solutions, hydrolyzing to formamides (not shown).

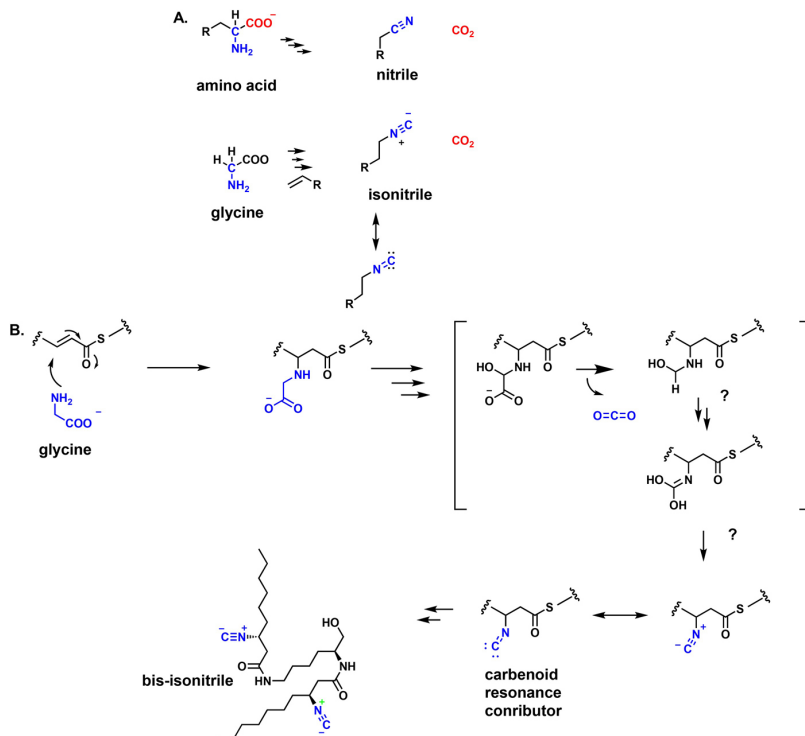


Figure 13.14 Biological formation of isonitriles. (A) Comparison of nitriles to isonitriles. (B) Pathway to isonitriles by addition of glycine amino group into an α,β -enoil-thioester.

In nitrile assembly from amino acids we have noted consecutive *N*-oxygenations of the amino group on the way to the *C*-nitroso and ultimately the cyano group with decarboxylation of the carboxylate group along the way. The original C–N bond of the amino acid persists in the triply bonded –CN product nitrile group.

A recent biosynthetic study of the formation of a bis-isonitrile lipopeptide natural product **24** from *Streptomyces coeruleorubidus* indicates that glycine is the precursor of the isonitrile moieties, again with decarboxylation of the C₁-carboxylate to CO₂. The key N–C bond-forming step in isonitrile group formation is attack of the deprotonated amine group of glycine on an α,β -enoil-S-carrier protein in a nonribosomal peptide synthetase (NRPS) assembly line (see Chapter 15) (Figure 13.14B). Subsequent decarboxylation is one oxidative step but a further iron-based desaturation (as yet uncharacterized) to create the triply bonded isonitrile is required in the biosynthetic pathway. The enzymatic *N*-oxidation strategies for nitrile and isonitrile formation by four-electron oxidation of amine starter units are fundamentally different and yield functionalities that are also differentially reactive.

13.8 Overview

This chapter has examined the chemical biology of cyanide formation (and briefly isocyanides) and utilization in bacterial and plant systems. Cyanide is an intriguing organic scaffold for nitrogen. The nitrogen atom is triply bonded in this simplest of

organic nitrogen compounds. Indeed, the cyanide ion is isoelectronic with N₂ but differs in being highly reactive as a ligand for iron, both in toxic (cytochrome C oxidase) and physiological contexts (bacterial hydrogenase ligand). Furthermore, the cyanide ion (about 1% abundance of total HCN at pH 7.2) is a rare kinetically stable biological carbanion source. The cyanide ion is highly polarized with the excess charge localized almost entirely on carbon. However, in the organic nitriles, such as pyrrolopyrimidines, the carbon atom in the nitrile group can be attacked by nucleophiles, such as cysteine thiolate side chains in nitrile group reduction, revealing chemical biology versatility to the cyano group.

More than 3000 plant species, including major crop plants, assemble and store cyanogenic glycosides as latent sources of toxic free cyanide ions. A subset of the common proteinogenic amino acids are subjected to iterative *N*-hydroxylation by an unusual class of cytochrome P450 oxygenases. Of a total of 200–300 plant-encoded cytochrome P450s, only the CYP79 and CYP71 subclasses catalyze *N*-hydroxylations. The *N,N*-dihydroxy metabolites, dehydrate to *C*-nitroso compounds followed by decarboxylation that isomerizes the nitroso group to an aldoxime. A separate P450 mediates *C*-hydroxylation and net oxidation of the oxime to the hydroxynitrile group. Alpha hydroxynitriles are carbonyl addition products where the cyanide carbanion can be viewed as the leaving group and slow decomposition to aldehyde and cyanide ion is observed. Therefore, plants convert these labile hydroxynitriles to *O*-glycosylated species which are indefinitely stable. Activation to release toxic cyanide in a real time defensive response means the plant has to carry out a glycosidase reaction and then action by a partner hydroxynitrilase. Two enzymatic steps lowers the probability of unwanted release of cyanide ions at the wrong time or place.

Separately, pathways to approximately three dozen pyrrolopyrimidine natural products, serving as nucleoside antimetabolites are built from the central nucleotide GTP. 7-Deazguanine scaffolds with 7-cyano substituents are often either pathway intermediates or end point metabolites. The organic nitriles are not precursors to toxic inorganic cyanide ion. The metabolism of such organic nitrile scaffolds involves ATP-fragmenting biosynthesis from carboxylates and ammonia to yield amides, then nitriles. In turn, the nitriles can be reduced by four-electron processes to the corresponding amine, tying together acyl amide biological chemistry and amine chemistry.

References

1. E. Antonini, M. Brunori, C. Greenwood, B. G. Malmstrom and G. C. Rotilio The interaction of cyanide with cytochrome oxidase, *Eur. J. Biochem.*, 1971, **23**, 396-400.
2. P. K. Yadav, K. Yamada, T. Chiku, M. Koutmos and R. Banerjee Structure and kinetic analysis of H₂S production by human mercaptopyruvate sulfur transferase, *J. Biol. Chem.*, 2013, **288**, 20002-20013.
3. P. Cressey and J. Reeve Metabolism of cyanogenic glycosides: a review, *Food Chem. Toxicol.*, 2019, **125**, 225-232.
4. U. G. Wganer, M. Hasslachner, H. Greiengl, H. Schwab and C. Kratky Mechanisms of cyanogenesis: the crystal structure of hydroxynitrile lyase from *Hevea brasiliensis*, *Structure*, 1996, **4**, 811-822.
5. R. A. Kahn, T. Farendorf, B. A. Halkier and B. L. Moller Substrate specificity of the cytochrome P450 enzymes CYP79A1 and CYP71E1 involved in the biosynthesis of the cyanogenic glycoside dhurrin in *Sorghum bicolor* (L.) moench, *Arch. Biochem. Biophys.*, 1999, **363**, 9-18.
6. T. Yeats, *Plantae; Plant Physiology News and Views*, 2018, cited 08 June 2020.
7. E. E. Conn Cyanogenic compounds, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1980,

8. R. M. Gleadow and B. Moller Cyanogenic glycosides: synthesis, physiology, and phenotypic plasticity, *Annu. Rev. Plant Biol.*, 2014, **65**, 155-185.
9. K. Gruber, G. Gartier, B. Krammer, H. Schwab and C. Kratky Reaction mechanism of the hydroxynitrile lyases of the alpha beta hydrolase superfamily, *J. Biol. Chem.*, 2004, **279**, 25001-25010.
10. T. N. Akopyan, A. Braunstein and E. V. Goryachenkova Beta-cyanoalanine synthase: purification and characterization, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 1617-1621.
11. L. Martinova Nitrile metabolism in fungi: a review of its key enzyme nitrilases with focus on their biotechnological impact, *Fungal Biol. Rev.*, 2019, **33**, 149-157.
12. R. M. McCarty and V. Bandarian Biosynthesis of pyrrolopyrimidines, *Bioorg. Chem.*, 2012, **43**, 15-25.
13. V. Bandarian and C. Drennan Radical-mediated ring contraction in the biosynthesis of 7-deazapurines, *Curr. Opin. Struct. Biol.*, 2015, **36**, 116-124.
14. M. T. Nelp and B. Vandarian A single enzyme transforms a carboxylic acid into a nitrile through an amide intermediate, *Angew. Chem., Int. Ed.*, 2015, **54**, 10627-10629.
15. V. M. Chikwana, B. Stec, B. W. K. Lee, V. de Crecy-Lagard, D. Iwata-Reuyl and M. A. Swairjo Structural basis of biological nitrile reduction, *J. Biol. Chem.*, 2012, **287**, 30560-30570.
16. R. P. Singh, G. Shelke, A. Kumar and P. H. Jha Biochemistry and genetics of ACC deaminase: a weapon to "stress ethylene" produced in plants, *Front. Microbiol.*, 2015, **6**, 937.
17. S. F. Yang and N. Hoffmann Ethylene biosynthesis and its regulation in higher plants, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1984, **35**, 155-189.
18. L. M. Mirica and J. Klinman The nature of O₂ activation by the ethylene-forming enzyme 1-aminocyclopropane-1-carboxylate oxidase, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1814-1819.
19. J. Lavill, C. Blumer, C. Von Schroetter, V. Gaia, G. Defago, C. Keel and D. Haas Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent *Pseudomonas fluorescens CHA0*, *J. Bacteriol.*, 1998, **180**, 3187-3196.
20. W. Lubitz, H. Ogata, O. Rudiger and E. Reijerse, *Hydrogenases. Chem. Rev.*, 2014, **114**, 4081-4148.
21. P. M. Vignais, B. Billoud and J. Meyer Classification and phylogeny of hydrogenases, *FEMS Microbiol. Rev.*, 2001, **25**, 455-501.
22. E. Pilet, Y. Nicolet, C. Mathevon, T. Douki, J. C. Fontecilla-Camps and M. Fontecave The role of the maturase HydG in the [FeFe] hydrogenase active site, *FEBS Lett.*, 2009, **583**, 506-511.
23. R. C. Driesener, M. Challand, S. E. McGlynn, E. M. Shepard, S. B. Eric, J. B. Broderick, J. W. Peters and P. L. Roach [Fe-Fe] hydrogenase cyanide ligands derived from S-adenosylmethionine-dependent cleavage of tyrosine, *Achie*, 2010, **49**, 1687-1690.
24. N. C. Harris, M. Sato, N. A. Hermann, F. Twigg, W. Cai, J. Liu, J. Downey, R. Khala, J. Martin, H. Koshino and W. Zhang Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptidase synthetases in actinobacteria, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**(27), 7025-7030.
25. A. Loulliset *al.*, The structure of a ferrous heme-nitro species in the binuclear heme a₃/CuB center of ba₃-cytochrome c oxidase as determined by resonance Raman spectroscopy, *Chem. Commun.*, 2015, **51**, 286-289.

One-electron Nitrogen Chemical Biology

14.1 One-electron Transfers To and From Inorganic Nitrogen Atoms

In the first section of the book on inorganic nitrogen metabolism in Chapters 2–5 much of the discussion centered on electron transfer processes. The metabolic transformations by microbes illustrate the full range of nitrogen oxidation states to span eight electrons, from $N = +5$ (nitrate) to $N = -3$ (ammonia). Two enzymes in different nitrogen reductive metabolic regimes, nitrogenase and assimilatory nitrite reductase, both carry out net six-electron reductions. In nitrogenase (FeMoCo) and also nitrite reductase (siroheme), the respective active sites contain iron atoms as the proximal electron donors to bound nitrogen substrates and intermediates. The iron-based redox cofactor identities are consistent with the six electrons being fed into either N_2 or nitrite ion one at a time traversing nitrogen oxidation states, 0 to -3 (both atoms of N_2 by nitrogenase) and from $+3$ to -3 (nitrite reductase), respectively. In turn, that electron transfer logic reinforces the one-electron redox switch between ferrous ion (Fe^{2+}) and ferric ions (Fe^{3+}). We will note transient higher oxidation states of iron (Fe^{4+}) and (Fe^{5+}) in certain enzyme microenvironments later in this chapter.

Bacterial denitrification of $2 NO_3^-$ ions to N_2 mandates five-electron reduction of each nitrate ion. The odd electron stoichiometry clearly involves at least one step involving a single electron transfer to a nitrogen atom. That catalyst is the denitrifying nitrite ion reductase. As noted in Chapters 4 and 5, denitrifying bacteria have evolved two types of nitrite reductases that carry out the one-electron reduction of nitrite ion to the nitric oxide radical, either by a di-iron center or a di-copper center (Figure 5.6). We have also noted in Chapter 5 that the radical species nitric oxide is a free intermediate in three of the inorganic nitrogen redox metabolic pathways: denitrification, anammox and nitrification. Much of the evolution of microbes using ammonia as an energy source or nitrate as a terminal electron sink developed around the use of heme and copper ions as one-electron transfer metals. Evidently, one-electron chemistry at nitrogen continues to be a staple of bacterial inorganic nitrogen metabolism 1:2 and a measure of the mechanistic accessibility and compatibility of nitrogen atoms as one-electron donors and acceptors in league with iron and copper catalytic centers.

Although the great bulk of eukaryotic nitrogen chemical biology appears to operate without redox change at nitrogen, the three-electron oxidative fragmentation of

arginine to citrulline and nitric oxide [3](#) (Chapter 12) likewise places nitric oxide radicals squarely in the realm of mammalian organic nitrogen redox metabolism, as emphasized in Section 14.3.5.

14.2 Two-electron vs. One-electron Redox Manifolds on Organic Nitrogen Scaffolds

By comparison with the redox ups and downs of inorganic nitrogen metabolites as anaerobic and aerobic microbes eke out a living chemolithoautotrophically, most of the organic nitrogen metabolic transformations occur at the $N = -3$ formal oxidation state. We have examined some of the many enzymatic examples of conversion of amines to amides, amides to amidines, amines to ureas and ureas to guanidines. These all occur without formal redox change at the reacting nitrogen centers ([Figure 14.1](#)). This also holds true for amines to peptide bonds in proteins, both by ribosomal and nonribosomal machinery, summarized in Chapter 15.

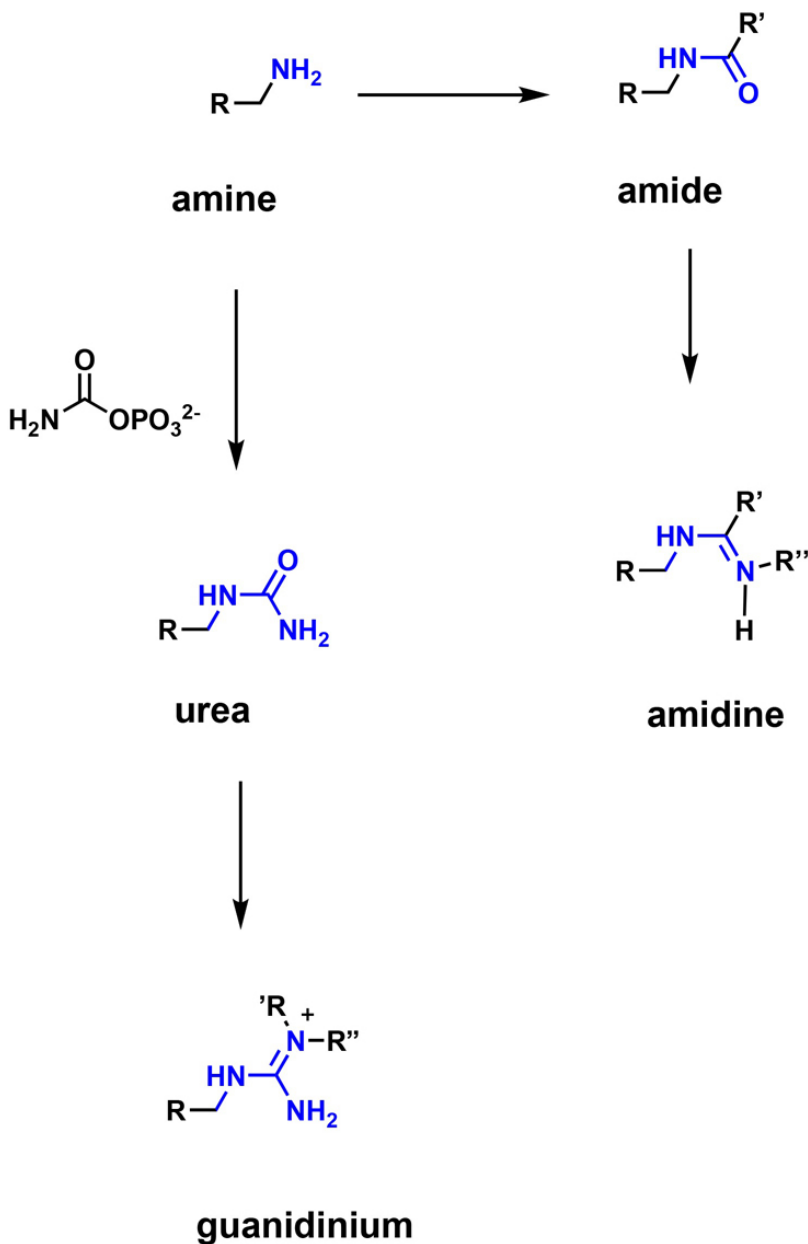


Figure 14.1 Organic nitrogen metabolism without redox change at nitrogen.

The reactions that fix ammonia into the organic frameworks of glutamate and glutamine likewise involve no change in nitrogen oxidation state. Nor do deaminases that liberate ammonia from amino acid scaffolds. The bioconstruction of pyrimidine and purine heterocyclic building blocks for RNAs and DNA are almost redox free. The one step in pyrimidine biosynthesis, catalyzed by dihydroorotate dehydrogenase, does not involve either of the two ring nitrogens (although flavins are part of the enzyme's redox inventory). The bicyclic ring framework of purines is assembled without redox

change. The tetrapyrrole biosynthetic pathway uses the chemistry of the pyrrole heterocycle without redox change at nitrogen. Also, we opened the door on pyrrole-based indole enamine chemical biology in Chapter 11 without oxidation state change at nitrogen.

In the above panoply of enzymatic transformation of organic nitrogen-containing scaffolds, we have focused on the basicity and nucleophilicity of nitrogens, largely amines and occasionally amide nitrogens, in terms of available lone pair on the particular nitrogen under consideration. All those bond-forming reactions where nitrogen acts as a nucleophile are in essence two-electron reactions, using both electrons in the lone pair.

The stability of organic amines and the host of other metabolites noted above in the presence of 21% O₂ in the atmosphere is due to mechanistic incompatibility for reaction. O₂, a ground state triplet, reacts extremely sluggishly with spin-paired molecules, allowing organisms, including humans, to exist in the thermodynamically favored oxidizing global atmosphere. Indeed, N₂ at 78% abundance and O₂ at 21% abundance coexist without detectable reactivity in earth's atmosphere. The favorable thermodynamics would send us all to CO₂ and H₂O at equilibrium.

Amines *can* react by one-electron rather than in two-electron reaction manifolds but the energy barriers are generally higher and reaction rates much slower. We look in the next sections at that two-electron–one-electron dichotomy in the context of O₂ reductive activation mechanisms by flavin-dependent oxidases and oxygenases vs. heme iron-based cytochrome P450 oxygenases.

14.2.1 PLP-dependent Transaminases are Redox Catalysts but Leave the Nitrogen Oxidation State Unchanged

In Chapter 6 we examined amino acid transaminases, as the most widespread enzymatic machinery in amino acid metabolism for moving amino groups between different α -ketoacid scaffolds. These are almost invariably pyridoxal phosphate (PLP)-dependent enzymes. They explicitly demonstrate that the B₆ coenzyme PLP can function as a two-electron redox coenzyme, as it shuttles between PLP–aldehyde and the two-electron reduced PNP–pyridoxamine phosphate oxidation state. The two-electron redox change is manifest in the coenzyme reductive half reaction by release of the keto acid product. Analogously, in the coenzyme reoxidative half reaction, the keto acid-2 is reductively aminated to amino acid-2 as the PNP form is oxidized by two electrons from the amine to the aldehyde PLP (Figure 14.2).

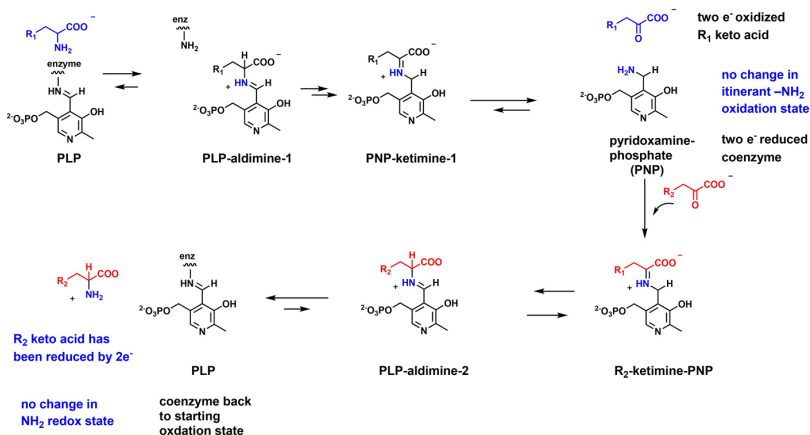


Figure 14.2 PLP-dependent transaminases: two-electron redox catalysts without change in redox status of the itinerant NH_2 group as it moves between keto acid scaffolds.

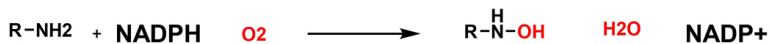
However, inspection of the amine group undergoing net transfer between the keto acid frameworks reveals no change in amine oxidation state. The NH_2 group of amino acid-1 becomes the amino group of PNP and then becomes the amino group of amino acid-2. There is no redox change at the NH_2 group being shuttled between all the keto acid partners in the cell.

In some analogy, as nicotinamide adenine dinucleotide ion (NAD^+)-dependent amino acid dehydrogenases convert the amino acids to ketoacids and ammonia as NADH is generated, the NH_3 in flight is unaltered in its formal nitrogen redox state. The substrate framework does undergo two-electron oxidation at C_2 as it goes from amine to ketone. The two electrons removed have to go somewhere and they always go into the pyridinium ring to generate the dihydronicotinamide NADH. The redox steps involving nicotinamide coenzymes are always obligatory two-electron steps, mediated by direct hydride transfer.

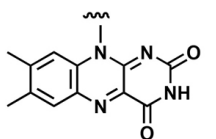
14.3 Organic Nitrogen Radical Metabolites

There are, however, several examples of enzyme-generated organic nitrogen radical intermediates, emphasizing that one-electron reaction manifolds are operant in organic as well as inorganic nitrogen metabolism. An obvious locus is in *N*-oxygenation metabolism. We have seen *N*-oxygenations of lysine and ornithine side chain amino groups in siderophore biosynthesis in Chapter 12. Also, in the biosynthesis of nitrile metabolites by thousands of plants, noted in Chapter 13, the initial steps are amino acid C_2 amine group oxygenations. Oxygen chemical biology is a good place to look for one-electron radical intermediates derived from cosubstrates, given the exclusive propensity of ground state O_2 to react by one-electron pathways.

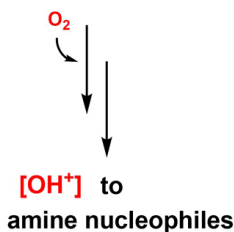
The oxygenations of N_6 of lysine and N_5 of ornithine in siderophore biosynthesis described in Chapter 12 are undertaken by metal-free enzyme catalysts. They harbor one equivalent of tightly bound flavin coenzyme, usually flavin adenine dinucleotide (FAD). In a different biological context, we noted the first steps in plant organonitrile metabolite assembly in thousands of plants begins with amino acid $\text{C}_2\text{-NH}_2$ *N*-oxygenations, but by a subclass of heme iron-based cytochrome P450 *N*-oxygenases. While both classes of oxygenases employ one-electron reaction manifolds in reductive activation of O_2 , they differ dramatically in two-electron vs. one-electron reaction manifolds in *the actual amine oxygenation step*. In turn, that difference reveals a fundamental distinction between $[\text{OH}^+]$ equivalent transfers by the FAD coenzymes vs. $[\text{OH}]$ transfer by the heme cofactor (Figure 14.3).



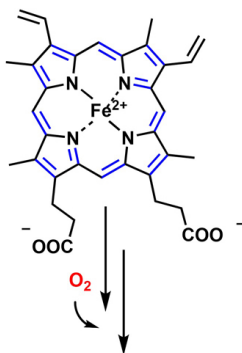
1. flavoenzyme oxygenases deliver $[\text{OH}^+]$ equivalents



flavin coenzymes



2. heme enzyme oxygenases deliver $[\text{OH} \cdot]$ equivalents



$[\text{OH} \cdot]$ to
amine radicals

Figure 14.3 *N*-oxygenation of substrate amines: NADPH as a common external donor of two electrons; flavoenzyme oxygenases transfer $[\text{OH}^+]$ equivalents while hemeprotein *N*-oxygenases transfer an $[\text{OH} \cdot]$ equivalent to amine radical forms of substrates.

14.3.1 Parallel One- and Two-electron Strategies in Substrate Amine *N*-Oxygenations

14.3.1.1 Inventory and Path of Electron Flow to FADH_2 or to Heme-Fe

Both the FAD-containing- and the heme-containing *N*-oxygenases carry out net four-electron reductions of O_2 , splitting it to H_2O and the *N*-OH-products. ^{4,8} Both classes use readily available reduced forms of the nucleotide coenzymes [reduced nicotinamide adenine dinucleotide phosphate (NADPH)] as external donors of two of those four electrons. The other two electrons come from the amine (NH_2) group undergoing oxidation by net two electrons to the *N*-hydroxy ($-\text{NHOH}$ = hydroxylamino) product. This is clear evidence for *N*-oxidation state change ($\text{N} = -3$ goes to $\text{N} = -1$). The issue is whether this occurs in a single two-electron step (flavins) or two one-electron steps (N at -3 to N at -2 to N at -1) (heme) *via* a nitrogen radical intermediate (N at -2).

Both classes of *N*-oxygenases use NADPH as a direct hydride transfer donor to reduce a tightly bound FAD by two electrons in one step. From there the paths of electrons diverge. In the FAD-based *N*-oxygenases we will note in the next paragraph that the FADH_2 is the reactive species for one-electron transfer to O_2 .

In the P450 monooxygenases the path for electrons to reach the active site heme group is longer and necessitates a two-electron–one-electron step down switch to accommodate the restriction that (ground state) Fe is an obligate one-electron redox metal. Typically, the electrons in the FADH_2 coenzyme will split and transfer one electron at a time to bound flavin mononucleotide (FMN), yielding a pair of

$$\text{R-NH}_2 \xrightarrow{\text{NADPH, O}_2} \text{R-N-OH} + \text{H}_2\text{O} + \text{NADP}^+$$

common initial hydride transfer

enz

2e⁻ in as hydride

enzyme

one e⁻ from FADH₂ to FMN

two 1e⁻ steps

[OH⁻] transfer

NADPH

FAD

FADH₂

FMN

FADH[•]

pair of flavin semiquinone radicals

flavoenzyme oxygenases: amine as nucleophile capturing [OH[•]]

FMNH[•]

1e⁻

Fe₄S₄ ox.

Fe₄S₄ red.

heme protein oxygenases amine radical coupling with [OH[•]]

[OH[•]] transfer to amine radicals

14.3.1.2 FADH₂-dependent Oxygenation and Transfer

In keeping with the requirement that O₂ reacts with organic compounds by one-electron routes, the FADH₂ form can donate one electron to molecular oxygen to generate a radical pair, the one-electron reduced superoxide anion radical and the FADH semiquinone radical. The ability of FADH₂ (and FMNH₂) to transfer one electron is due to the fact that the resultant odd electron flavin semiquinone (Figure 14.5A) has the odd electron delocalized over several atoms in the isoalloxazine tricycle, including at N₅ and N₁. This is nitrogen radical chemistry practiced by hundreds of flavoenzyme monooxygenases (regardless of specific substrate oxygenated). 7:9

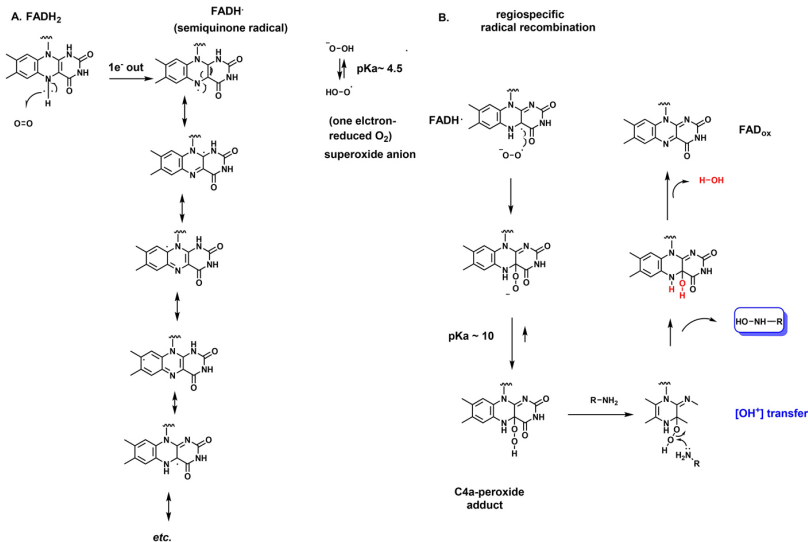


Figure 14.5 (A) The reduced form of flavins (such as FADH₂) can react rapidly with O₂ for one-electron transfer because the resultant flavin semiquinone radical (one-electron oxidized) is kinetically accessible due to the delocalization *via* many resonance contributors to the radical structure. (B) Radical recombination of the superoxide anion radical and the flavin semiquinone radical is directed to the bridgehead C4a position to give the covalent flavin-4a-OOH peroxy adduct. This acts as a donor of an electrophilic oxygen equivalent to nucleophilic amine cosubstrates.

The superoxide anion radical and the [FADH[•]] pair, in the close confines of the active site then undergo radical recombination, directed to form a regiospecific C–O bond as the covalent bridgehead flavin-hydroperoxide (Figure 14.5B). That bridgehead carbon is the C_{4a} carbon of the flavin: so this is a flavin-4a-OOH. At this juncture, O₂ has undergone a net two-electron reduction (in two 1e⁻ steps) to the oxidation state of an alkyl peroxide.

At this point in the reaction cycle the amine cosubstrate gets involved. The flavin C4a–O–OH adduct has a weak O–O single bond, typical of peroxides and that can be fragmented by attack of the amine cosubstrate. As shown in Figure 14.5B, the amine typically attacks the terminal (distal) oxygen of the alkyl peroxide. The result is transfer of the OH to the amine to form the N–OH product. The mechanism is that the lone pair on the amine nitrogen has attacked the distal oxygen in a *two-electron manifold*. The OH group in flight is transferred as an [OH[•]] equivalent.

The two electrons that were in the O–O bond of the 4a-OOH adduct remain with the proximal oxygen and protonate to give the flavin-4a-OH adduct. Inspection reveals that this is the water addition product to oxidized FAD. Thus, facile loss of water across N₅–H, C_{4a}–OH yields back oxidized FAD and coproduct HOH, completing the catalytic cycle (Figure 14.5B).

In sum, once the enzyme–FADH₂ complex has been formed, there is radical chemistry for one-electron reduction to O₂ and subsequent radical recombination to the C4a-OOH covalent adduct. The [FADH[•]] radical intermediate has unpaired electron density at both N₅ from which [H] departed and at N₁, nitrogen radical chemistry. However, the actual oxygen transfer to cosubstrate amine nitrogen is a two-electron reaction manifold. No radicals in this part of the reaction coordinate.

14.3.2 One-electron Transfer in All Phases of Cytochrome P450 N-oxygenations

It is instructive to compare that above riboflavin coenzyme-based two-electron oxygenation logic with the strategy for oxygen transfer implemented by the Cyp79 subclass of plant P450 enzymes (some 3000 plants use this strategy). 7 Figure 14.6 serves as a reminder that heme protein oxygenases can accept a first and then a second electron to get to a ferric hydroperoxy state, written as two-electron reduced O_2 (Fe^{III} -OOH) as the top axial ligand. One common fate of that species is use of two electrons from the Fe^{III} to cleave the coordinated weak O-O bond – logically analogous to the FAD-4a-OOH oxygen-oxygen bond cleavage?

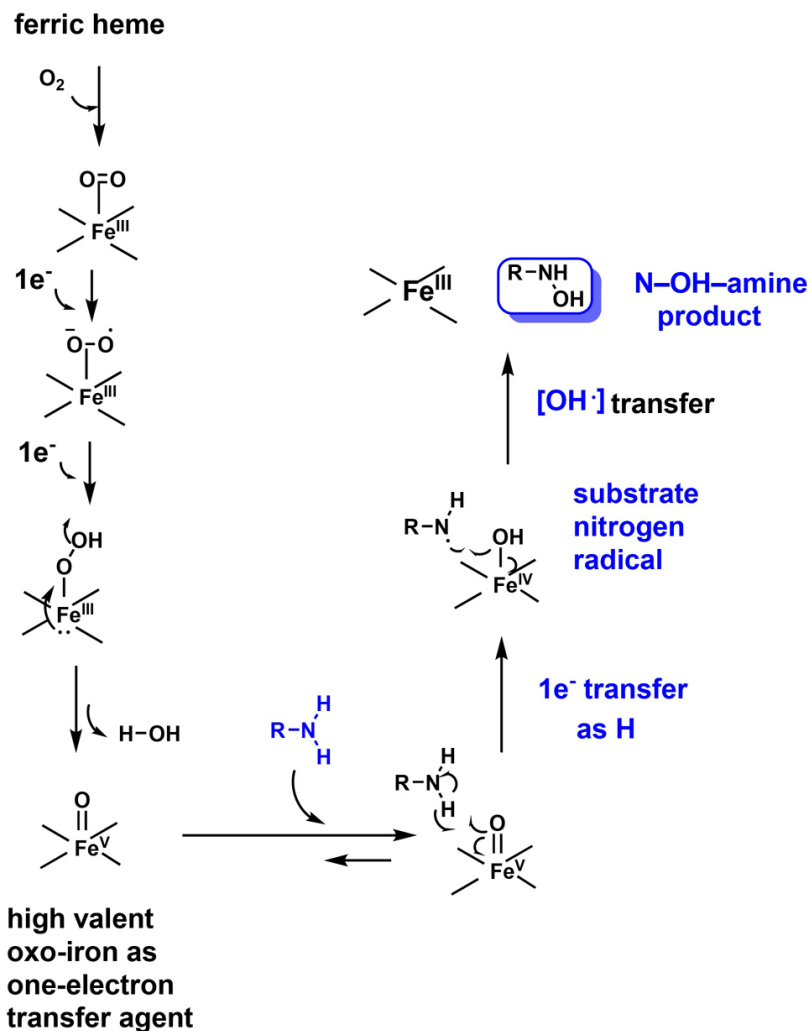


Figure 14.6 Hemeprotein cytochrome P450 monooxygenases proceed via $Fe^V=O$ high valent iron oxides species that are robust oxidants and initiate catalysis by hydrogen atom abstraction from substrate amines to generate amine radicals that

react with [OH] transfer.

That act generates a high valent heme $\text{Fe}^{\text{V}}=\text{O}$ oxoiron as a robust oxidant, so robust that it can cleave unactivated C–H or N–H bonds. Here, the substrate amine N–H gives up a hydrogen atom [H]. That creates a formal $\text{Fe}^{\text{IV}}\text{--OH}$ species close to the substrate amine radical ($N = -2$). Oxygen transfer can be viewed as a radical rebound, transfer of [OH] to the substrate amine radical. ¹⁰ This forms the N–O bond in the nascent N-hydroxy product ($N = -1$). Because this is a homolytic transfer of [OH], the other electron in the $\text{Fe}^{\text{IV}}\text{--OH}$ bond flows back into iron to give the Fe^{III} state (Figure 14.6).

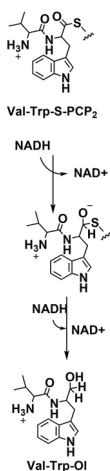
The fundamental logic of iron-based heme-type oxygenases is to generate high valent oxoiron species as powerful oxidants, powerful enough to “rip” a hydrogen atom away from a C–H or N–H bond in a docked substrate as prelude to radical [OH] transfer to the substrate radical. Members of the Cyp79 family of plant N-oxygenases, repeat the radical oxygenation process a second time, as summarized in Figure 13.4, to give the *N,N*-dihydroxy nascent product. That can lose water nonenzymatically to produce the *nitroso* acid product ($N = +1$). Such C-nitroso carboxylates can undergo facile decarboxylation driven in part by isomerization of a C=N bond to an N=O double bond, specifically the rearrangement of a C-nitroso to an aldoxime group. The decarboxylation has left the original sp^3 hybridized carbon of the amine substrate in the sp^2 hybridization state of an aldehyde in oxime linkage.

Typically, one thinks of the high valent oxoiron species in heme iron enzymes as the stronger “biological oxygenating guns” than the flavin-4a-OOH groups. ¹⁰ Flavin-linked enzymes catalyze oxygen transfer to electron rich substrates that can capture the $[\text{OH}]^+$ equivalents from the Flavin 4a-OOH (organic peroxide). Iron-based oxygenases generate much more reactive and oxidizing high valent oxoiron intermediates that can oxygenate “unactivated” substrates. Most often those are C–H bonds at unactivated carbon centers. In the Cyp79 case the unusual substrate class are the primary amines of amino acids. Given equivalent amine starting substrates, the FAD-based oxygenases and the heme-based Cyp79 oxygenases, reveal the parallel two-electron–one-electron strategies as the choice of $[\text{OH}]^+$ - vs. $[\text{OH}]^-$ -like transition states, and which cofactor provides the more robust oxygen transfer oxidant.

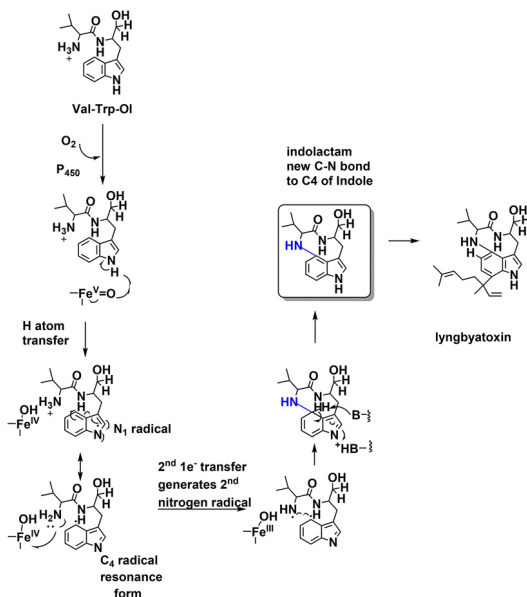
14.3.3 Nitrogen Radical Coupling to Carbon and Nitrene Transfer by Cytochrome P450s

We examine the bacterial biosynthesis of the active chemical in swimmer's itch, lyngbyatoxin and its core scaffold indolactam V to highlight the physiological accessibility of organic nitrogen radicals ¹¹ that *do not suffer N–O bond formation* in the universe of cytochrome P450 oxygenases. An O_2 -dependent cytochrome P450 enzyme generates a pair of nitrogen radicals, an indole N_1 radical and a valyl N_1 radical that then couple in C–N bond formation in the bacterial biosynthetic pathway to lyngbyatoxin (Figure 14.7). That toxin acts as a phospholipase C agonist in humans. The core of lyngbyatoxin is a nine-membered ring, known as indolactam V, arising from that internal C–N bond formation in the dipeptide alcohol L-valyl-L-tryptophanol. ¹² That dipeptidyl alcohol in turn is a product of a two-module nonribosomal peptide synthetase (NRPS) assembly line with a four-electron reductive release of the enzyme-tethered Val–Trp–acyl–S–peptidyl carrier protein domain (Figure 14.7A).

A. Two Module NRPS
with 4e⁻ reductive release



B. Indolactam Formation
by a P450 generating a
pair of Nitrogen Radicals



C. Bezastatins

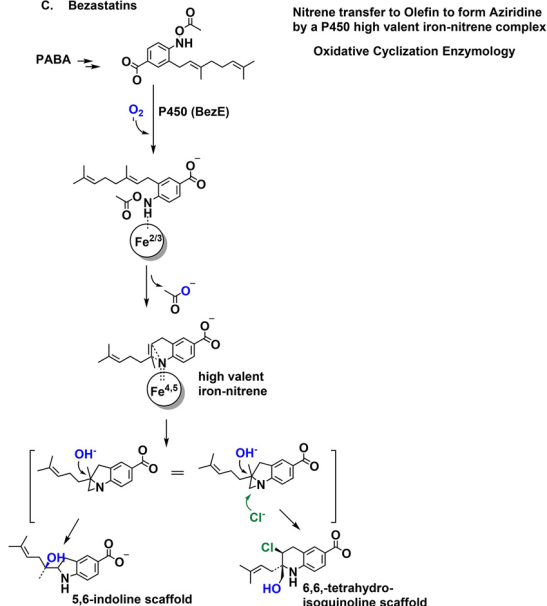


Figure 14.7 Cytochrome P450s that lead to N–C bonds instead of N–OH bonds. (A) Formation of L-valyl–L-tryptophanol by a two module nonribosomal peptide synthetase (NRPS) with a reductase terminal domain. (B) C–N bond formation by an O₂-dependent cytochrome P450: initial Indole N₁ radical formation by the Fe^V=O oxidant has partial unpaired electron density at C₄. Valyl–N₁ radical generation by electron transfer to Fe^{IV}–OH sets the stage for radical coupling and indolactam V nine member ring formation. Subsequent prenylation at C₇ yields

lyngbyatoxin. (C) Formation of two bezastatin isomers by the P450 BezE is proposed to involve an $\text{Fe}^{\text{V}}=\text{N}$ nitrene on the way to a nascent aziridine that can be opened to indole or tetrahydroisoquinoline scaffolds. Adapted from [ref. 14](#) with permission from American Chemical Society, Copyright 2018.

The formation of the indolactam Val-N₁-Trp C₄-indole N-C bond occurs by a variant of canonical heme protein oxygenase logic. [11](#) O₂ is reductively activated and split to the $\text{Fe}^{\text{V}}=\text{O}$ oxidant. One-electron transfer of a hydrogen atom from substrate gives the $\text{Fe}^{\text{IV}}-\text{OH}$ species that acts normally as donor of an [OH] equivalent in a rebound radical coupling. However, two alternate steps intervene. There is no oxygenation of substrate. This P450 qualifies as a “thwarted oxygenase”. [13](#)

First, the hydrogen atom transfer is proposed to come from the N₁-H of the indole ring of the tryptophanol moiety to create the N₁ indole radical. As shown in [Figure 14.7B](#) that has a resonance contributor with unpaired electron density at C₄. Second, the N₁-NH₂ group of the valyl moiety then undergoes a one-electron transfer to the $\text{Fe}^{\text{IV}}-\text{OH}$, deactivating it as the $\text{Fe}^{\text{III}}-\text{OH}$ species and ultimately releasing that OH ligand as the second molecule of water from the starting O₂.

Meanwhile, that electron transfer has generated the second substrate nitrogen radical in the catalytic cycle, the valyl-N₁-NH species. [11](#) That nitrogen radical recombines with the C₄ radical resonance contributor on the indole ring to form the new C-N bond, thereby constructing the unusual nine-membered ring in the indolactam V framework. A subsequent prenylation at C₇ of that indole ring by geranylgeranyl-pyrophosphate acting as the C₃ carbocation isomer completes biosynthesis of the lyngbyatoxin molecule.

There are two points to note in lyngbyatoxin formation. First, is that the indolactam V formation involves sequential generation of a pair of substrate nitrogen radicals, emphasizing one-electron reaction manifolds, even though the regiospecific radical recombination is to C₄, not to form a new N-N bond. Second, while indolactam ring formation shows accessible one-electron delocalization chemistry of indole rings, the subsequent geranylation at C₇ of that indole ring reveals the two-electron delocalization chemistry of indoles to put partial negative charge density at multiple sites on the benzene ring component (indoles as radical or carbanions for physiological aminations *versus* alkylations). The regiochemistry of N-C bond formation at C₄ and then C-C bond formation at C₇ in turn illustrate the orientational selectivity imposed by binding forces at the two enzymes' active sites.

One additional variant of a cytochrome P450 carrying out nitrogen oxidative enzymology is found in the formation of two distinct bezastatin isomers. [14](#) [Figure 14.7C](#) shows that BazE is also an O₂-dependent member of the cytochrome P450 superfamily. Like the P450 above, it generates high valent oxoiron species but does not successfully transfer an oxygen atom to cosubstrate. Instead, the proximal oxidant is proposed to involve not the prototypic high valent $\text{Fe}^{\text{V}}=\text{O}$ species, but instead an $\text{Fe}^{\text{V}}=\text{N}$ (*iron-nitrene*) oxidant. A reactive three-member aziridine ring would then be generated as a bound intermediate that can be opened by water or a chloride ion to give the 5,6 or 6,6-fused indoline or tetrahydroisoquinoline scaffolds, respectively.

14.3.4 Cyanogenesis Involves Nitrogen Radicals in Cyanide Ion Formation

Nitrogen radicals are involved in two of the enzymes discussed in Chapter 13 on cyanogenesis. The cyanide ion and carbon monoxide ligands required as ligands for iron centers in both the bis-iron hydrogenases and the nickel-iron hydrogenases ([Figure 13.13](#)) derive from controlled fragmentation of L-tyrosine. The relevant enzyme has an Fe₄-S₄ cluster with a coordinated S-adenosylmethionine (SAM). On one-

electron input from an external donor (such as a reduced ferredoxin protein), the ligated SAM undergoes homolysis of the adenosyl-S⁺ bond to liberate the 5'-deoxyadenosyl radical within the enzyme active site. [15](#)

We have commented in Chapter 11 that adenosyl-B₁₂ is the other route by which homolytic release of [dA] in the middle of a catalytic cycle unleashes a robust radical oxidant. The [dA] abstracts a hydrogen atom from the tyrosine cosubstrate, creating the *amine radical* as the first step in the ultimate fragmentation of the tyrosine scaffold into *para*-cresol, carbon monoxide and a cyanide ion (Figure 13.12). Homolysis of the SAM-coordinated to Fe₄-S₄ clusters to generate [dA] as a radical initiator turns out to be the more general biological solution in perhaps hundreds of thousands of enzymes than the adenosyl-B₁₂ machinery, for which there are less than two dozen known cases. [10](#)

The second example of *in situ* generation of a nitrogen radical on an amino acid framework occurs in the production of ethylene by plants as a fruit ripening hormone (Chapter 13). The cyanide ion is a stoichiometric coproduct with ethylene as the unusual cyclopropane ring-containing 1-amino-cyclopropane-1-carboxylate undergoes fragmentation. In that case a radical reaction manifold is initiated by a nonheme high valent iron in an oxoiron adduct rather than the deoxyadenosyl radical, a reminder of nature's two sources of radical generators in metabolism. The resultant nitrogen radical on the amine of ACC in turn initiates the fragmentation of the strained cyclopropane ring and the net oxidative release of ethylene gas as the desired hormone product. [16](#) It seems that cyanide ion comes along as a byproduct of the ethylene-releasing cyclopropane ring fragmentation. These cyanide molecules are used as nucleophiles by *beta*-cyanoalanine synthase enzymes to detoxify cyanide before it can cause problems by ligation to heme and nonheme iron centers.

14.3.5 Nitric Oxide Synthase Revisited for One-electron Transfer Mechanisms

One can divide the reaction mechanism of nitric oxide synthase isoforms into two steps, each involving reductive activation of O₂ at the heme site. [17](#) The first two-electron oxidation and formation of an N-O bond yields bound *N*-hydroxyarginine (Figure 14.8A) by a classic P450 type monooxygenase: an *amine substrate radical* by [H] abstraction and then transfer of an [OH] equivalent in a radical rebound to form a covalent -N-OH bond (N = -1). High valent (Fe^V) iron oxo species are generated as the powerful oxidants for H atom abstraction from bound substrates. In the nitric oxide synthase (NOS) cases that would be a nitrogen radical on the guanidine side chain followed by [OH] equivalent rebound.



It is during the consumption of the second molecule of O₂ that the nitric oxide synthases display unusual redox stoichiometry. Presumably, the chain of events up to forming a high valent heme iron oxide is similar to the first oxygen transfer cycle. Although the mechanism of this second catalytic cycle is not understood in full detail, a proposal by Groves and Wang [17](#) suggested that in this second catalytic cycle, homolysis of the axial coordination bond between the nascent N-OH-arginine and Fe^{III} would yield Fe^{II} and the *N*-hydroxyarginine radical that also can be written with unpaired electron density at the carbon atom between the three guanidine nitrogens

(Figure 14.8A). Then, O₂ binding, followed by a single-electron transfer in from the FAD–FMN–heme relay system could yield the indicated heme–Fe^{II}–peroxy bridged N–oxide adduct (with a new C–O bond) from the indicated N–OH arginine radical resonance form connecting with a superoxide ion equivalent. That adduct could fragment to release the nitric oxide radical, citrulline with the ureido oxygen deriving from O₂, and a high valent Fe^{IV}=O heme species. This is the species that could be captured by one-electron transfer from tetrahydrobiopterin to get back to an Fe^{III} heme oxidation state.

This fragmentation of N–OH–arginine to nitric oxide and citrulline is, in part, the reverse of the reaction in denitrifying bacteria where nitrite is reduced by one electron to nitric oxide (Chapters 4 and 5). In those bacteria two strategies evolved for that one electron reduction. One was to employ heme C and heme-D cofactors where heme D is the site for axial coordination of a nitrite ion followed by one-electron input from an obligate one-electron heme donor. The other evolved solution involves copper^{II}–copper^I redox cycles for the net one-electron transfer to coordinated nitrite ion.

The structure of the human inducible NOS is shown as a dimer in Figure 14.8B (PDB file 3NOS). A close up view of that structure reveals two cofactors in the oxygenase catalytic domain, heme and tetrahydrobiopterin. (The flavins are in a separate domain that binds the external reductant NADH and thus not displayed in this domain.) Also bound in the structure in green is N–hydroxyarginine, the product of the first oxygenation event and the substrate for the second oxygenation–fragmentation that yields citrulline and nitric oxide radical.

14.4 Nitric Oxide Biology

Nitric oxide biology can be split into its roles in inorganic vs. organic nitrogen metabolism. Bacterial inorganic nitrogen metabolism in both anaerobes and aerobes involves the nitric oxide radical as a freely diffusing metabolite, exiting a producing enzyme active site with lifetimes long enough to encounter the next enzyme in the pathway, be it a reductase (in denitrifiers) or an oxidase (in nitrifiers). It is not clear if the nitric oxide radical plays any signaling roles in microbial biology.

In contrast, nitric oxide arises in mammalian tissues by only a single enzyme pathway (Figure 14.9), catalyzed in humans by two constitutive and one inducible nitric oxide synthases. Unlike bacterial inorganic nitrogen metabolism, nitric oxide is not an intermediate in a redox–electron transport pathway. Instead, it is a terminal metabolite radical, with specialized signaling properties by virtue of nitrosyl radical–ferrous heme complex formation. We turn to that mammalian biology next.

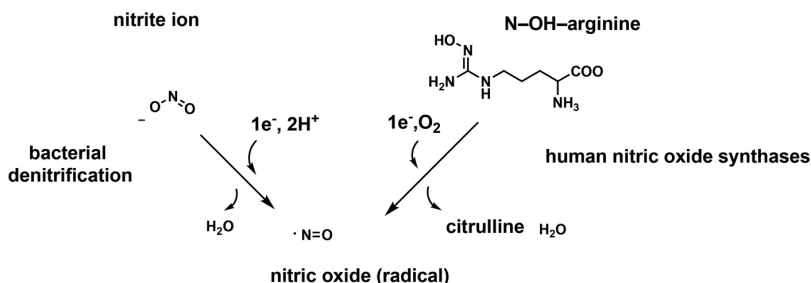


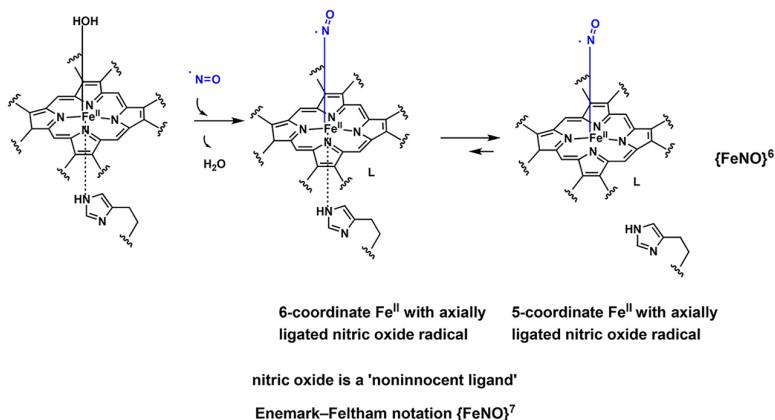
Figure 14.9 Two biological routes to the nitric oxide anion: inorganic nitrogen redox metabolism generates nitric oxide from nitrite ions by *one-electron reduction*.

Mammalian nitric oxide synthases generate nitric oxide from *N*-hydroxyarginine by *one-electron oxidation*.

14.4.1 Direct and Indirect Effects of Nitric Oxide in Higher Eukaryotes

Nitric oxide, made by endothelial, inducible or neuronal NOS (eNOS, iNOS or nNOS) in different human tissues, is both water soluble and neutral (uncharged) so can partition into and through membrane spaces. The physiological signaling role for [NO], at nanomolar concentrations, is to bind with highest affinity (approximately 10^{-9} M) as the top axial ligand to the ferrous iron in the regulatory heme group of guanylate cyclase 3·18·19 (Figure 14.10A). That binding sets off electronic and conformational changes first at the heme, for example dissociation of the bottom axial histidine's imidazole ring, and then propagation of the conformational change to the cyclic GMP (cGMP) synthase active site, resulting in up to a 300-fold increase in rate of formation of 3,5'-cGMP, 20 a second messenger in cell signaling (Figure 14.10B).

A. ferrous heme–nitrosyl complex formation



B. catalytic efficiency enhanced 300-fold by nitric oxide binding as top axial heme ligand at regulatory site

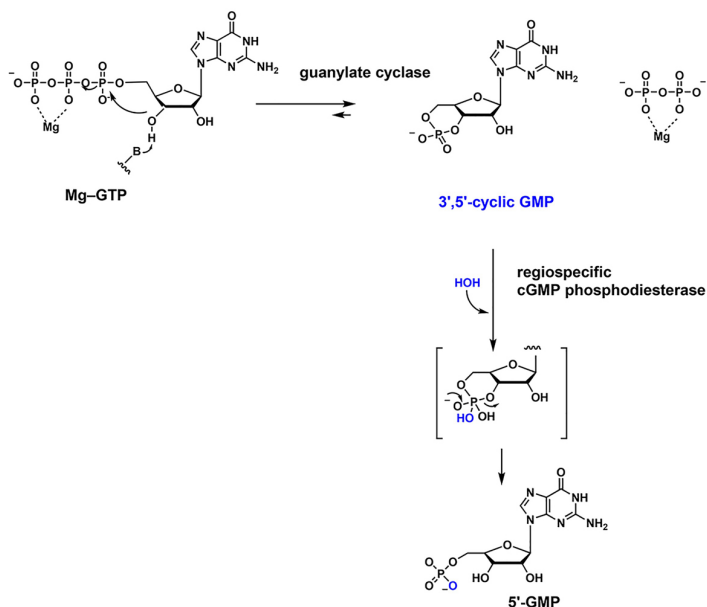


Figure 14.10 (A) The nitric oxide radical has a high affinity for binding to ferrous heme as a top axial nitroxyl radical ligand. Because that nitroxyl radical is a non-innocent ligand with regard to electron interaction with the tetrapyrrole iron, the system has been described as an $\{\text{FeNO}\}^7$ complex to recognize the interaction and the inability to specify exactly where the unpaired electron density from the nitric oxide radical resides. (B) The guanylate cyclases that are the physiological targets of nitric oxide have a regulatory heme-b site that becomes occupied by the nitrosyl radical as in part A. That ligation is communicated allosterically to the guanylate cyclase site to increase the catalytic constant (k_{cat}) up to 300-fold, transiently elevating the regulatory second messenger molecule 3',5'-cyclicGMP.

The interaction of NO with ferrous heme is some 10^4 -fold tighter than that with

ferric heme. One can write the resultant iron nitrosyl complex as the nitric oxide radical coordinated to Fe^{II} , but it has been known for decades that nitric oxide is a “non-innocent ligand” (Figure 14.10A). For example, if there is partial transfer of the nitric oxide unpaired electron to Fe^{II} , that would generate an Fe^{I} -nitrosyl cation (NO^+) contributor to the complex, however small the contribution. A common shorthand for formulating the existence of an electronic interaction is the notation of Enemark-Feltham (E-F) counting the number of π electrons in the complex. **21** Ferrous iron brings six, nitric oxide brings one unpaired electron so in the E-F notation one writes the ferrous heme-nitrosyl as a $(\text{FeNO})^7$ complex, **21** implicitly recognizing there is non-innocent interaction by not specifying where the nitric oxide π electron resides (Figure 14.10A). A ferric heme-NO complex would be in this notation an $(\text{FeNO})^6$ complex.

As local nitric oxide concentrations transiently rise into micromolar concentrations as a result of NOS activities, most of this higher burst interacts with red cell oxygenated hemoglobin **22** (Figure 14.11). The molecular oxygen- $[\text{N}=\text{O}]$ complex reacts and rearranges to a nitrite ion, a major detoxification removal of the nitric oxide radical. The iron^{II} in hemoglobin is concomitantly oxidized to iron^{III} . This is methemoglobin, requiring reduction by the flavoenzyme methemoglobin reductase to regain its capacity as an iron^{II} reversible carrier of O_2 .

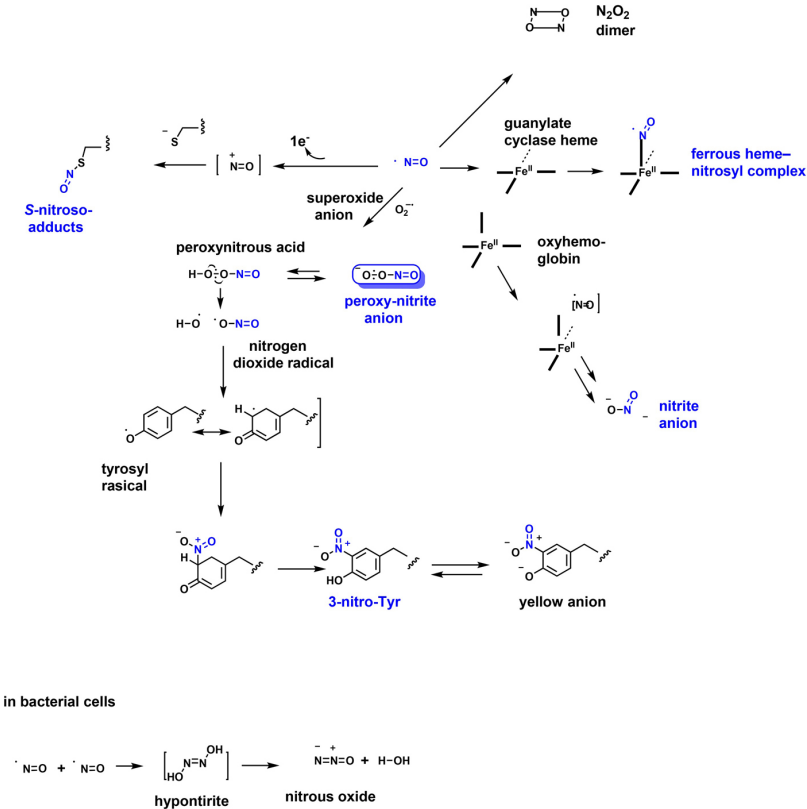


Figure 14.11 Multiple reactions of nitric oxide radicals in mammalian cells. Direct effects at low concentration include binding to the regulatory heme-b of guanylate cyclases. At higher transient levels of $[\text{NO}]$, reaction with superoxide to form peroxynitrite, with ferrous hemoglobin to oxidize it to ferric methemoglobin, or

with thiols to *S*-nitroso adducts, lead to secondary reactions.

Figure 14.11 also shows that nitric oxide radicals can be quenched by superoxide ion, generated by a variety of reactive oxygen stress routes and by leakage of superoxide ion from adventitious one-electron reductions of O_2 by the mitochondrial electron transport chain. The radical recombination yields the peroxynitrite ion. ²³ The protonated form is peroxynitrous acid. With a pK_a of 6.3, both the protonated peroxynitrous acid and deprotonated peroxynitrite ion are present. The neutral peroxynitrous acid can fragment homolytically to two problematic radicals that can cause damage, the hydroxyl radical [OH] and the nitrogen dioxide radical [NO₂].

One of the telltale outcomes from nitric oxide arises from the indirect reaction by the granddaughter metabolite (the nitrogen dioxide radical) [NO₂] with tyrosyl side chain phenoxyl radicals. Those phenoxyl radicals would or could arise during reactive oxygen species (ROS = superoxide, peroxide, hydroxyl radical) stress. The combination of the tyrosyl radical and nitrogen dioxide radical yields 3-nitrotyrosyl residues in proteins (Figure 14.11). The nitro group electronically lowers the pK_a of the phenolic-OH group from approximately 10.5 to approximately 7.5. The nitrophenol anion absorbs in the approximately 420 nm region and imparts a detectable yellow color to nitrotyrosyl residues in proteins. ¹⁸

A final result of transiently elevated nitric oxide production is the nonenzymatic formation of *S*-nitroso adducts with glutathione or cysteinyl thiolate side chains in proteins (Figure 14.11). In turn, the nitroso groups are transferable to other thiolate anions for redistribution to the most accessible and/or nucleophilic thiolate side chains with some attendant functional consequences. It is not clear if the wandering nitric oxide radicals are first oxidized by one electron to the nitrosonium cations (NO⁺) that are then captured by thiolate anions or whether there are sufficiently high levels of one-electron oxidized thiol radicals to react directly with the nitric oxide for radical recombination to the observed *S*-nitrosothiol adducts. In sum, while the highest affinity target in cells for nitric oxide is the heme group of guanylate cyclase, transiently elevated levels above nanomolar concentrations set off a variety of (unwanted?) reactions that are sequelae to nitrogen radicals on the loose.

Thomas *et al.* ¹⁸ have noted that the manifold effects of nitric oxide can be defined along two axes. One axis is to separate out direct effects of [NO] vs. indirect effects that arise from reaction or interception of [NO] by cellular molecules. The second axis is to rank order effects of [NO] based on the concentrations of [NO] at which the effects occur.

For example, Thomas *et al.* ¹⁸ cite that the enhancing effects on cyclic-GMP metabolism occur at concentrations of [NO] as low as 1–30 nanomolar. These are clearly physiological as the dissociation constant (K_d) for NO binding to the heme on its prime target protein guanylate cyclase is approximately 1 nanomolar. Transient phosphorylation of the kinase Akt is mediated at 30–100 nM concentrations of [NO]; stabilization of the hypoxia transcription factor HIF-1 α occurs at 100–300 nM; p53 phosphorylation is observed at >400 nM. Nitrosative stress occurs at (transient) concentrations of [NO] greater than 1 micromolar. Readers wishing to delve further into the physiological and/or pathophysiological relevance of those [NO] concentration regimes are referred to the article by Thomas *et al.* ¹⁸ as an entry point to the literature.

14.4.2 Interaction with Guanylate Cyclase

Guanylate cyclases carry out the intramolecular attack of the 3'-OH of the ribose moiety of Mg-ATP on the adjacent P α , an internal nucleotidyl transfer releasing PPi and creating the fused 5–6 ring junction in 3',5'-cyclic GMP (Figure 14.10B). Along

with the congeneric 3',5'-cyclic AMP (formation catalyzed by adenylyl cyclases), cGMP and cAMP are the preeminent small molecule second messengers in cells. Cyclic-GMP activates a range of target proteins, such as cGMP-dependent protein kinases and also various ion channels for a broad range of signal transduction effects. The duration and intensity of cGMP signaling is modulated by a set of cGMP phosphodiesterases at different intracellular loci to create a moment-to-moment balance of [cGMP] (Figure 14.10B). The intersection of [NO] and cyclic GMP has been a burgeoning topic in cell signaling and physiology for the past three decades and is beyond the limited focus of this chapter.

Soluble guanylate cyclases are typically heterodimeric mixtures of isoforms of α - and β - subunits. Key to their regulation is the presence of a tightly bound heme group that is five coordinate Fe^{II} in the resting form. 20-24 The top axial position is open, while a histidine imidazole side chain occupies the bottom axial ligand. That heme group is in essence a sensor for nitric oxide. NO binds and forms a top axial nitrosyl radical-iron complex (Figure 14.10A). The bottom histidine can dissociate to give a mixture of nitrosyl heme geometries that drive conformational changes to activate the cyclase activity. Mg-GTP is converted to cGMP and PPi and the downstream signaling effects of elevated [cGMP] kick in. These encompass a wide range of tissue responses, including smooth muscle relaxation (lowering of blood pressure), inhibition of platelet aggregation and synaptic plasticity. The precise molecular mechanisms for transmission of nitrosyl radical-heme occupancy to the catalytic center for activation of Mg-GTP intramolecular cyclization has remained unresolved due to an absence of X-ray and electron microscopy structures of this conformationally malleable regulatory enzyme system.

14.5 Overview

Nitrogen atoms in biology can be redox-active and traverse eight electrons from the most reduced forms (NH_3 , RNH_2 , $\text{N} = -3$) to the most oxidized form as nitrate ions (NO_3^- , $\text{N} = +5$). Redox changes are the main province of inorganic nitrogen chemical biology, as evidenced by the six-electron traverses in nitrogenase and in assimilatory nitrite reductase catalytic cycles. Most of organic nitrogen metabolism is nonredox, with amine nitrogens predominating as reactants.

However, there are several instances where the oxidation state of nitrogen attached to organic scaffolds undergoes enzyme catalyzed changes. Among the most salient examples are nitric oxide formation from arginine, amino acid *N*-oxygenations and in cyanogenesis. The amino acid *N*-oxygenations allow comparison of the two-electron logic used by flavin coenzymes vs. the one-electron logic of cytochrome P450 monooxygenases in catalyzing the oxygen transfers to substrate amine groups. The separate fragmentations of tyrosine and of aminocyclopropane carboxylate to cyanide ions likewise illustrate the different logic of $\text{Fe}_4\text{-S}_4\text{-SAM}$ enzymes vs. nonheme iron oxygenases to create *in situ* oxidants that generate nitrogen radicals to initiate scaffold fragmentations. Those catalytic cycles convert a singly bonded $-\text{C}-\text{NH}_2$ to the triply bonded free $-\text{CN}$ ion.

References

1. J. D. Caranto and K. Lancaster Nitric oxide is an obligate nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 817-822.
2. Y. Shiro, H. Sugimoto, T. Tosha, S. Nagano and T. Hino Structural basis for nitrous

- oxide generation by bacterial nitric oxide reductases, *Philos. Trans. R. Soc., B*, 2012, **367**, 1195-1203.
3. U. Forstermann and W. Sessa Nitric oxide synthases: *regulation and function*, *Eur. Heart J.*, 2012, **33**, 829-837.
4. S. Eswaramoorthy, J. Bonanno, S. K. Burley and S. Swaminathan Mechanism of action of a flavin-containing monooxygenase, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9832-9837.
5. J. Olucha and A. Lamb Mechanistic and structural studies of the *N*-hydroxylating flavin monooxygenases, *Bioorg. Chem.*, 2011, **39**, 171-177.
6. C. Mugge, T. Heine, A. G. Baraibar, W. J. H. van Berkel, C. E. Paul and D. Tischler Flavin-dependent *N*-hydroxylating enzymes: distribution and application, *Appl. Microbiol. Biotechnol.*, 2020, **104**, 6481-6499.
7. R. A. Kahn, T. Farendorf, B. A. Halkier and B. L. Moller Substrate specificity of the cytochrome P450 enzymes CYP79A1 and CYP71E1 involved in the biosynthesis of the cyanogenic glycoside dhurrin in *Sorghum bicolor* (L.) *moench*, *Arch. Biochem. Biophys.*, 1999, **363**, 9-18.
8. S. Imrmisch, A. McCormick, G. A. Boeckler, A. Schmidt, M. Reichelt, B. Schneider, K. Block, J.-P. Schnitzler, J. Gershenzon, S. B. Unsicker and T. B. Kollner Two herbivore-induced cytochrome P450 enzymes CYP79D6 and CYP79D7 catalyze the formation of volatile aldoximes involved in poplar defense, *Plant Cell*, 2013, **25**, 4737-4754.
9. C. T. Walsh and T. Wenciewicz Flavoenzymes: versatile catalysts in biosynthetic pathways, *Nat. Prod. Rep.*, 2013, **30**, 175-200.
10. C. T. Walsh and Y. Tang, *Biosynthesis of Natural Products*, Royal Society of Chemistry Press, London UK, 2017.
11. I. Abe, I. Morita and T. Mori Enzymatic formation of indolactam scaffold by C-N bond-forming cytochrome P450 oxidases in teleocidin biosynthesis, *Chemistry*, 2020, **27**(9), 2963-2972.
12. J. A. Read and C. T. Walsh The lyngbyatoxin biosynthetic assembly line: chain release by four-electron reduction of a dipeptidyl thioester to the corresponding alcohol, *J. Am. Chem. Soc.*, 2007, **129**(51), 15762-15763.
13. C. T. Walsh and B. S. Moore Enzymatic Cascade Reactions in Biosynthesis, *Angew. Chem., Int. Ed.*, 2019, **58**(21), 6846-6879.
14. H. Tsutsumi, Y. Katsuyama, M. Izumikawa, M. Takagi, M. Fujie and N. Satoh *et al.*, Unprecedented Cyclization Catalyzed by a Cytochrome P450 in Benzastatin Biosynthesis, *J. Am. Chem. Soc.*, 2018, **140**(21), 6631-6639.
15. J. B. Broderick, B. Duffus, K. S. Duschene and E. M. Shepard Radical S-adenosylmethionine enzymes, *Chem. Rev.*, 2014, **114**, 4229-4317.
16. L. M. Mirica and J. Klinman The nature of O₂ activation by the ethylene-forming enzyme 1-aminocyclopropane-1-carboxylate oxidase, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1814-1819.
17. J. T. Groves and C. C. Wang Nitric oxide synthase: mechanisms and models, *Curr. Opin. Chem. Biol.*, 2000, **4**, 686-695.
18. D. D. Thomas, L. Ridnour, J. S. Isenberg, W. Flores-Santana, C. H. Switzer, S. Donzellie, P. Hussain, C. Vecoli, N. Paolocci, S. Ambs, C. Colton, C. Harris, D. D. Roberts and D. A. Wink The chemical biology of nitric oxide. Implications in cellular signaling, *Free Radical Biol. Med.*, 2008, **45**, 18-31.
19. A. J. Gow The biological chemistry of nitric oxide as it pertains to the extrapulmonary effects of inhaled nitric oxide, *Proc. Am. Thorac. Soc.*, 2006, **3**, 1501-1552.
20. B. G. Horst, A. Yokom, D. J. Rosenberg, K. L. Morris, M. Hammel, J. H. Hurley and M. A. Marletta Allosteric activation of the nitric oxide receptor soluble guanylate cyclase by cryo-electron microscopy, *eLife*, 2019, **8**, e50634.

21. R. D. Harcourt Construction of valence bond structure for $[\text{FeNO}]_7$ nitrosylheme and non-heme complexes, *Nitric Oxide*, 2017, **69**, 51-55.
22. A. J. Gow and J. S. Stammler Reactions between nitric oxide and hemoglobin under physiological conditions, *Nature*, 1998, **391**, 169-173.
23. R. Radi Oxygen radicals, nitric oxide and peroxynitrite: redox pathways in molecular medicine, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 5839-5848.
24. M. Russwurm and D. Koesling NO activation of guanylate cyclase, *EMBO J.*, 2004, **23**, 4443-4450.

Section IV

In a sense, the final two chapters, comprising Section IV, are an effort to steer a path into and through the enormous subject of nitrogen proteomics by remaining focused on the two aspects of the underlying chemistry of nitrogen functional groups in proteins. The first aspect is the peptide bond, the sole covalent linkage between the amino acid building blocks in intracellular proteins (the disulfide bonds of proteins in oxidizing microenvironments are reduced as dithiols within reducing compartments of cells). The second aspect is the chemistry available to different nitrogen side chains of seven out of twenty proteinogenic amino acids when incorporated into proteins.

This approach is orthogonal to most treatments of those two topics: the first dealing with protein biosynthesis and the second topic on posttranslational modifications (PTMs). For example, in the universe of almost 200 different types of posttranslational modifications, this can be only a partial look at the scope of such modifications, leaving out the serine, threonine and tyrosine phosphorylations that number in the hundreds of thousands of human proteins. Also, the *S*-prenylations and palmitoylations are not included. However, the interested reader can readily dip into the mammoth PTM, protein phosphorylation and protein lipidation current literature elsewhere. The categorization of types of reactions of the nitrogen side chains in proteomes presented here by underlying chemistry is not so easy to find elsewhere.

Chapter 15 takes up peptide bond chemistry in three distinct biological contexts, comparing the chemical logic and the resultant constraints on the types of activated aminoacyl groups employed biologically. Peptide bonds are just a subclass of amide bonds and the same energetic and mechanistic requirements set forth in Chapter 6 (Section III) apply. For each amide or peptide bond formed, the carboxylate partner must be converted from a potential oxygen nucleophile to a carbonyl group electrophile.

ATP is the universal cellular phosphoryl transfer currency to be spent to drive otherwise unfavorable equilibria forward. Amide synthases of all kinds direct the substrate carboxylates to attack either the outside or the inside phosphoric anhydride in the ATP side chain to create aminoacyl phosphates or aminoacyl-AMPs. That is the first branch point. The amide synthases making aminoacyl phosphates that get captured by nucleophilic amine components are often members of the ATP grasp superfamily. The amide or peptide bond is accompanied by the telltale cleavage pattern of ATP to ADP and Pi. This is *prima facie* evidence that these are nonribosomal processes.

The second split comes among the amide synthases that make tightly bound aminoacyl-AMP intermediates. All eukaryotic protein biosynthesis uses the logic of capture of the aminoacyl groups by cognate tRNAs to make the *oxoester* linkage in the charged aminoacyl tRNAs. Once ferried to the ribosome by protein chaperones, these oxoesters are converted into amide or peptide bonds by the ribosomal transferase center, acting as an RNA catalyst. The order of the 20 proteinogenic amino acids incorporated is dictated by the coding sequence of a given messenger RNA through

triplet base pairing with the anticodon region of the tRNA moiety.

The other alternative, used by some bacterial and fungal enzymes, is to build nonribosomal peptide synthase (NRPS) assembly lines. The key chemical difference is formation of amino acyl *thioesters* instead of oxoesters, retaining more of the thermodynamic activation of the labile aminoacyl-AMPs but with enhanced lability to hydrolysis compared with the oxoesters in freely diffusing aminoacyl tRNAs. Thus, the aminoacyl thioesters in NRPS assembly lines are covalently tethered to peptidyl carrier protein domains. The order of amino acid incorporation is not controlled by an mRNA sequence but rather by the order and specificity of the adenylation domains in the NRPS assembly line. Consequently, a broad range of nonproteinogenic amino acid monomers are incorporated into nonribosomal peptides.

The ribosomal and nonribosomal strategies converge with amide or peptide bond formation. The driving force is oxoester to amide at the ribosome vs. thioester to amide in the NRPS assembly line. In both nanomachineries, the growing peptidyl chain is sequestered and intermediates are not released. Chain termination is exclusively hydrolytic at the ribosome but can be hydrolytic, reductive or macrocyclizing in the thioesterase domains at the end of NRPS assembly lines.

Once proteinogenic amino acid monomers have been installed in peptide linkage and the amino and carboxy groups tied down in those consecutive amide bonds, only the side chains are available for reaction. The six nitrogen-containing side chains in proteomes are the imidazole heterocycle, the N_6 -amine of lysine, the guanidinium group of arginine, the amides of asparagine and glutamine and the indole ring of tryptophan residues. Of these the imidazole ring nitrogens of histidine and the primary side chain of lysine are anticipated to be potentially functional as nucleophiles. The guanidinium and amides seem at the outset to be unlikely nucleophilic candidates, but we will examine the few cases where those expectations are contradicted. The indole group of free tryptophan, tryptamine and tryptophan-containing diketopiperazines shows off its enamine chemistry in many plant- and fungal-based natural products, but is notably suppressed, with one exception, in proteomes.

Imidazole, with its pK_a at 6.0, offers a higher fraction of deprotonated nitrogen than lysine (pK_a approximately 10) and arginine (pK_a approximately 12.5) and functions in different enzymes as either a general base in reaction with protons or as a nucleophile, in reaction with phosphorus in ATP to form $N\text{-PO}_3^{2-}$ phosphoramidate intermediates.

Lysine side chains, when deprotonated, behave as one would expect for a nucleophilic primary amine, undergoing imine formations, *N*-alkylations and *N*-acylations. When the *N*-alkylations and acylations are posttranslational modifications, as in histone H3 and H4 tails, there are opposing or restoring dealkylation and deacylation enzymes to terminate the covalent signaling functions.

Arginine side chains have a rich chemical life in protein structures. The predominant cationic guanidinium form is an effective bidentate catalyst for electrostatic interactions with glutamic acid and aspartic acid side chain carboxylates. The arginine guanidinium group also excels in electrostatic interaction with posttranslationally-generated phosphoserine, phosphothreonine and phosphotyrosine side chains. Less obvious are *N*-phosphorylation of arginine in bacterial proteins, marking them for proteolysis, monomethylation and demethylation of the guanidinium side chain in histones and other proteins, hydrolytic deamination to citrulline side chains that can be neoantigens and ADP-ribosylations.

The amide side chains of asparagine and glutamine are not as unreactive as one might anticipate. Notably, thousands of asparagine amide nitrogens are glycosylated on proteins passing through eukaryotic cell secretory compartments on the way to cell surfaces or secretion into the extracellular milieu. Each of these require the asparagine residue carboxamide nitrogen to behave, albeit weakly, as a nitrogen nucleophile. Glutamine carboxamides are usually quietly stable in proteins but transglutaminases

can seek out specific glutamine side chains on substrate proteins, carry out a glutaminase half reaction to release ammonia and form the protein-Glu-S-transglutaminase covalent intermediate. That adduct can be deacylated by a lysine side chain from a second protein, generating an interprotein amide crosslink. This is the nexus of glutamine and lysine posttranslational amide formation.

Finally, the sole known expression of the indole ring propensity to form carbanionic transition states at several carbon sites around the indole ring is found in the C₂-mannosylation of a tryptophan protein residue. One can propose a mechanistic variant of the known enamine chemistry of indoles for this unusual posttranslational C-glycosylation.

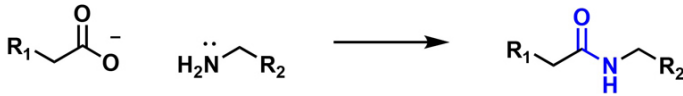
Peptide Bond Formation Strategies in Nonribosomal and Ribosomal Assembly Lines

15.1 Peptide Bonds

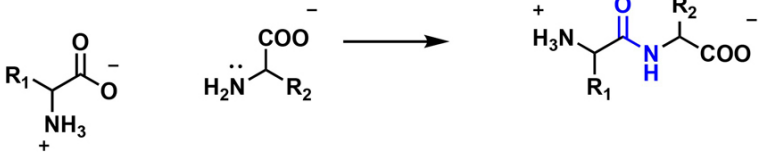
Peptide bonds, formed between two amino acid reactants, are a subclass of amide bonds ([Figure 15.1](#)). They are the backbone of all proteins and most peptides in organisms. Peptides are smaller than proteins, sometimes digestive fragments thereof, sometimes built on separate peptide assembly lines as discussed in Section 15.5. Also indicated in [Figure 15.1](#) is the requirement for energy input to drive the uphill formation of peptide or amide bonds from the amine and carboxylate components. Cells typically spend at least one ATP for each peptide bond formed.

typically requires
consumption of ATP
to activate each
carboxylate component

amide formation



peptide formation



**proteins are linear strings of
amino acid building blocks
covalently connected by peptide bonds**

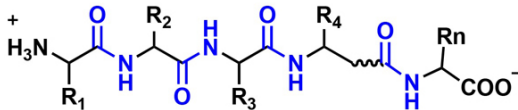


Figure 15.1 Peptide bonds are a subclass of amide bonds where both reactants are amino acids. ATP hydrolysis is required to activate the carboxylate partner and drive equilibrium in the direction of amide or peptide bond accumulation.

In terms of sheer numbers, peptide bonds in proteins dwarf every other type of cellular amide bond. In a mammalian cell with up to 25 000 separately encoded proteins in the genome, if the average protein size is approximately 300 amino acid residues (approximately 33 kilodaltons), then to produce a single copy of every protein would require 7.5 million amino acid residues. Protein copy number varies widely from a few per cell, say 10 at a lower limit to 10^5 copies of metabolically central proteins. If there were an average of 1000 copies of cell of the average sized protein that would yield approximately 7.5 billion peptide bonds in a cellular proteome. Multiply by 4×10^{13} estimated human cells in the body and one arrives at

approximately 3×10^{23} peptide bonds in a human. This represents dedicated peptide-amide bond enzymology.

15.2 Strategies for Amino Acid Monomer Activation for Peptide Bond Formation

Given the enormous commitment of cells and organisms to peptide bond biosynthesis, it is worth reexamining the mechanistic strategies, energy costs and differential activation of various chemical types of common aminoacyl groups in metabolism. As noted in [Figure 15.2](#), all the common α -amino acids, including the 20 proteinogenic ones, but also such key nonproteinogenic metabolites, such as ornithine, citrulline and homocysteine, are predominantly zwitterionic under physiologic conditions. The ammonium to amine pK_a values cluster between pH 9 and pH 10, while the carboxy groups dissociate to the resonance-stabilized carboxylate anions with pK_a values in the range of pH 4 to pH 5. Thus, neither the ammonium group cations nor the carboxylate anions are in the right ionization states for low energy amide or peptide bond formations between pairs.

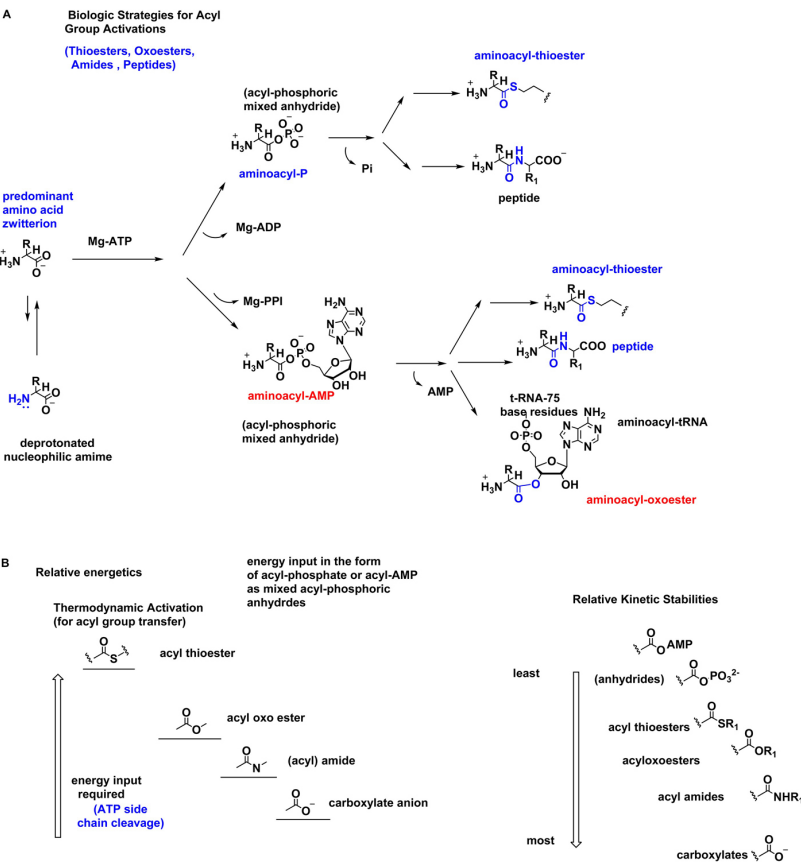


Figure 15.2 Biological strategies for acyl group activations. (A) Amino acids to aminoacyl-phosphoric anhydrides. The aminoacyl phosphates can go on to amides

or peptides (ATP grasp enzymes) or to aminoacyl thioesters. The corresponding aminoacyl-AMPs can go on to thioesters, oxoesters and eventually peptide bonds. (B) The relative energetics and kinetic stabilities of aminoacyl derivatives in biological systems.

15.2.1 Conversion of Ammonium Cations to Amine Nucleophiles by Deprotonation

The deprotonation of ammonium cations to the minor amine species seems to be a general property of enzyme catalysts that use amines either as bases for reaction with protons, or as nucleophiles for reaction with electrophiles, such as carbonyl carbons. Indeed, in all the previous chapters where organic amine nitrogens act as nucleophiles, we have sometimes explicitly noted the minor amine species as a proximal reactant, or often assumed that a given enzyme could accumulate a sufficient amount of the amine form to catalyze a given reaction. The activation of carboxylate ions, converting them from the anions that are potentially nucleophilic, not electrophilic, to electrophilic species suitable for capture by amine nucleophiles has required a covalent activation strategy, as we explicitly noted in Chapter 6.

15.2.2 Carboxylate Activation via Mixed Acyl-Phosphoric Anhydrides at the Expense of ATP Side Chain Cleavage

Every time an amide or peptide bond is formed, both in soluble reactants leading to a single amide bond formation, or in ribosomal or nonribosomal assembly lines, where tens to hundreds to thousands of peptide bonds are formed iteratively before full length or growing chain release, *an ATP cosubstrate molecule is cleaved*. Two patterns of cleavage at the phosphoric anhydride side chain are observed. ¹

Attack by the substrate (amino acid in this case) carboxylate can be directed to either the terminal P_{γ} or the internal P_{α} of Mg-ATP, depending on the orientation imposed on bound substrates by a given enzyme type. Attack at P_{γ} results in phosphoryl transfer to the carboxylate oxyanion and release of ADP. The phosphorylated carboxyl of an amino acid (shown in the upper part of [Figure 15.2A](#)) is termed an acyl phosphate. Chemically, it is an anhydride between a carboxylic acid and inorganic phosphoric acid. Such anhydrides are thermodynamically activated but generally kinetically rather labile in physiological milieus.

The alternative regiochemistry of ATP side chain phosphoric anhydride bond cleavage, at P_{α} , generates the aminoacyl-AMP species and releases inorganic pyrophosphate (PPi) (shown in the bottom half of [Figure 15.2A](#)). Aminoacyl-AMPs are also mixed acyl-phosphoric anhydrides, equally activated thermodynamically for acyl group transfer and kinetically as labile, if not more so, than the corresponding acyl phosphates. Aminoacyl-AMP mixed anhydrides are typically so tightly bound by the enzymes that form them that they almost never leave the active sites. ^{2,3} Otherwise, they would rapidly hydrolyze in bulk solution and dissipate the thermodynamic activation, wasting an ATP.

15.2.3 Harvesting the Thermodynamic Activation of Mixed Acyl-Phosphoric Anhydrides in Kinetically Stable Acyl Derivatives

Enzyme-catalyzed formation of both acyl phosphates and acyl-AMPs have achieved the goal of thermodynamic activation of carboxylate groups ([Figure 15.2B](#)). However, these mixed anhydrides are too labile kinetically to unwanted hydrolysis to serve as

reliable diffusing packets of acyl group transfers, e.g. for the approximately 10^{10} peptide bonds that humans make in every cell cycle. The harvesting of the aminoacyl phosphates is typically by capture of the activated acyl or aminoacyl groups by oxygen or thiolate anion nucleophiles.

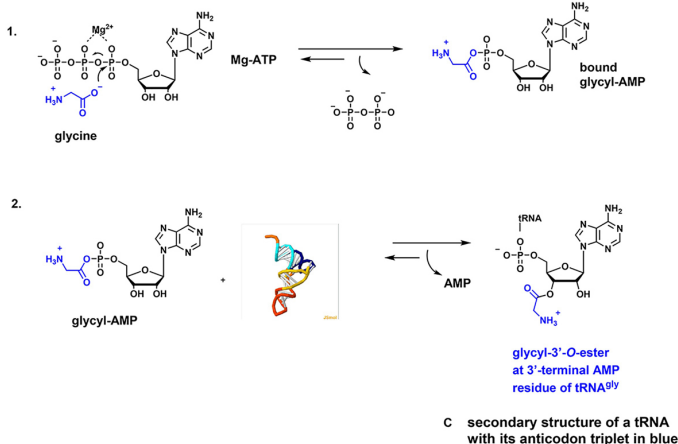
For acyl phosphates coenzyme A is the preferred thiol cosubstrate. ¹ The resultant acyl coenzyme As (CoAs) maintain the acyl group thermodynamic transfer potential but now have stabilities of several hours, appropriate to serve as diffusible packets of acyl group transfer cellular currencies (Figure 15.2A). Acyl-AMPs, such as acetyl-AMP, are likewise capturable by CoASH in the active site of enzymes to form the same acetyl CoA thioester as the above acetyl-phosphate pathway.

However, in particular for the proteinogenic 20 amino acids activated as aminoacyl AMPs by their cognate aminoacyl-tRNA synthases, the aminoacyl groups are transferred to *oxygen nucleophiles not sulfur nucleophiles* (Figure 15.2A). ² The products are aminoacyl oxoesters, less thermodynamically activated than aminoacyl thioesters (Figure 15.2B) but still retaining enough thermodynamic acyl group transfer potential to be the reagents for all peptide bonds made at the peptidyl transferase center of ribosomes.

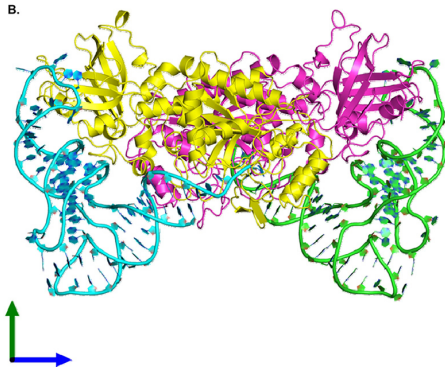
The oxygen nucleophiles in the second half reaction of aminoacyl tRNA synthetases (Figure 15.3A) are the cognate transfer RNAs (Figure 15.3B). The tRNA population consists of molecules with 64 different trinucleotide anticodon sequences, typically located around positions 35–40 in the approximately 75 residue long tRNAs. ^{4,5} (The amino acid oxo esters in charged tRNAs are at the 3'-terminal CCA trinucleotide, out at residues 73–75.) In bacteria 61 of the anticodons code for one of the twenty amino acid building blocks while 3 of the 64 are chain termination codons. All the anticodon triplets base pair with mRNA codons at the ribosome to determine which of the 20 amino acid building blocks will be inserted as the peptidyl chain grows. A typical tRNA structure is shown in Figure 15.3C, with the anticodon but also with the terminal CCA trinucleotide denoted (e.g. at residues 73–75 of the tRNA). It is the terminal adenosine residue A₇₅ that functions as an oxygen nucleophile by way of the 3'OH of its ribose moiety in the second half of the amino acyl tRNA synthetase reaction cycle (these are some 2'-O-acylations but the 2' and 3'-O-aminoacyl groups can interconvert). The aminoacyl transfer from the AMP mixed aminoacyl phosphate to the aminoacyl-oxoester linkage in aminoacyl tRNA is favored thermodynamically (Figure 15.2) and proceeds to completion.

A Aminoacyl- tRNA synthetase
two half reactions

Example: glycyl tRNA synthetase



B.



yeast aspartyl tRNA synthetase
cocrySTALLIZED with tRNA^{ASP}

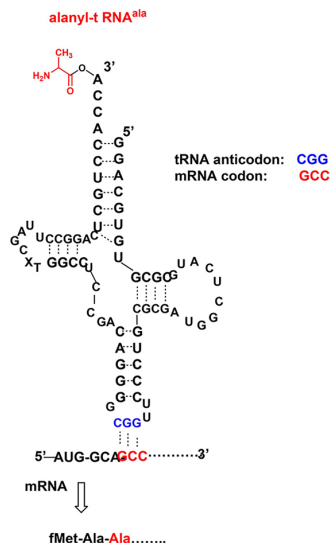


Figure 15.3 (A) Two-step reaction mechanism for aminoacyl-tRNA synthetases. First half reaction: selection of amino acid (glycine in this case) and activation as the aminoacyl-AMP, tightly bound. Second half reaction: transfer of the activated aminoacyl group (electrophilic at the carbonyl carbon) to the 3'-OH of the C-terminal AMP moiety at the 3'-end of a cognate glycyl tRNA to form the aminoacyl-O-tRNA (oxoester) ready to be chaperoned to the ribosome. (B). Three-dimensional structure of a canonical transfer RNA. PDB1asy bi.ac.uk/pdbe/entry/pdb/1asy. (C) Two-dimensional structure of a canonical transfer RNA, noting the C-C-A trinucleotide at the 3'-end where the aminoacyl group is appended in oxoester linkage to the 3'-OH of the terminal AMP residue. Also shown is the location of the anticodon triplet far in space from the aminoacyl group in loaded tRNAs.

Each aminoacyl-tRNA charged with its cognate aminoacyl moiety in oxoester linkage is kinetically stable enough to pair with a chaperone protein that escorts it to the peptidyl transferase center of the ribosome. We will return to the peptidyl chain initiation, elongation and termination phases of ribosomal machinery in the context of peptide bond growth after examining the ATP grasp family (ATP to ADP and Pi, not to AMP and PPi).

15.3 ATP Grasp Family of ADP-forming Amide Bond Synthetases

A family of ATP-cleaving, acyl transfer enzymes termed the ATP grasp family was initially categorized by bioinformatic analysis [6](#) in the late 1990s. The initial set of enzymes had 21 members but the numbers have grown several-fold since that initial classification almost 25 years ago. Even so, the ADP-forming ATP grasp enzyme family is less than 20% of the size of the protein kinase family in human proteomes.

The ATP grasp designation derives from structural information that the enzymes typically have two domains that coalesce to surround and grasp the ATP substrate. [7](#) ([Figure 15.4](#)). A unifying mechanistic feature is that all the ATP grasp enzymes make acyl phosphate intermediates, arising from attack of a cosubstrate carboxylate on the terminal P_γ in the side chain of ATP. The resultant electrophilic carboxylic phosphoric anhydride can be captured on the carbonyl carbon by a nucleophilic atom of the substrate. Most often these are amine nucleophiles, with the result of an amide bond in the products. When both cosubstrate pairs are amino acids, the product amide is in fact a peptide bond.

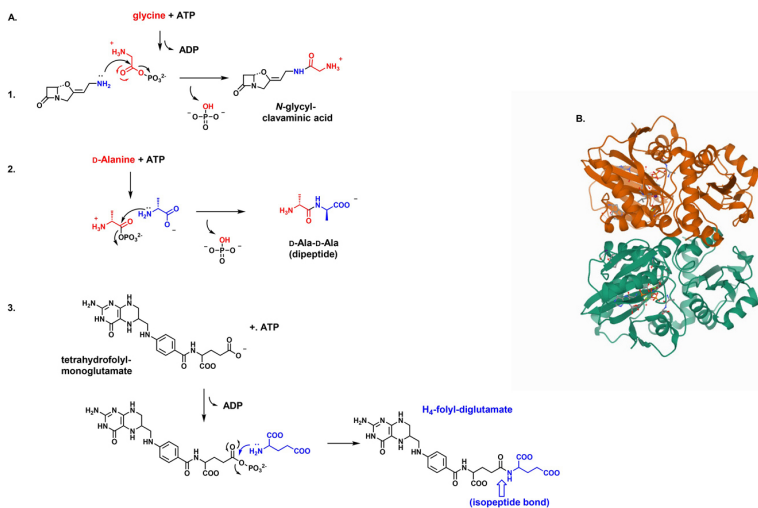


Figure 15.4 ATP grasp enzymes. (A) Three ATP grasp enzymes that use ATP to generate acyl phosphate intermediates on the way to amide, peptide or isopeptide bond formations. (B) Typical three-dimensional fold of an ATP grasp enzyme illustrated by the structure of D-alanine–D-alanine ligase from the bacterium *Thermus thermophilus*. (PDB file 2YDG).

This is the logic of the two enzymes that convert glutamate (Glu), cysteine (Cys) and glycine (Gly) into the abundant thiol-containing tripeptide glutathione. Both the γ-Glu–

Cys synthetase and the subsequent glutathione synthetase are ATP grasp enzymes, making the anticipated acyl phosphate precursors to the two amide bonds.

Analogously, the same logic applies for the three enzymes of [Figure 15.4A](#). The first enzyme in that figure is involved in the microbial metabolic pathway for the beta-lactam antibiotic clavamate. The acyl donor is glycine, activated by ATP cleavage as glycyl phosphate and captured by the side chain amine to give the isopeptide bond in glycyl-*N*-clavamate. [8](#) The second entry is the bacterial enzyme that provides dipeptide D-alanyl-D-alanine required for crosslinking of cell wall peptidoglycan to provide mechanical strength to that osmotic barrier feature of the cell wall. As shown, the D-alanine indicated in red is activated as the aminoacyl-phosphate while the D-alanine depicted in blue acts as a nitrogen nucleophile. The X-ray structure of the thermophilic bacterial D-Ala- D-Ala ligase from *Thermus thermophilus* [9](#) is shown in [Figure 15.4B](#).

The third example in [Figure 15.4A](#) deals with mammalian metabolism of ingested folic acid, vitamin B₉. Two metabolic events convert the oxidized folyl-monoglutamate to the active coenzyme forms. [10](#) First, as discussed in Chapter 10, is the four-electron reduction of the right hand ring portion of the bicyclic pterin ring to the tetrahydro pyrazine form in which the bottom nitrogen is nucleophilic. [11](#) Second, is the fact that the human enzymes utilizing coenzyme forms of vitamin B₉ recognize three to seven oligoglutamyl chains attached to the *para*-aminobenzoyl moiety. Otherwise, the folate forms are inactive for lack of binding recognition. The glutamyl-adding enzyme is shown converting the H₄-folyl monoglutamate to the H₄-folyl diglutamate. That will be repeated one to five more times to produce the active oligoglutamyl forms of folate coenzymes (see Chapter 10 for their roles in one-carbon metabolism).

The glutamyl-adding enzyme mechanism reveals two points. One is that the acyl phosphate component here is the H₄-folyl monoglutamate, to provide the electrophilic glutamyl C₅ carboxyl for amide bond formation *via* capture by the free amine of cosubstrate glutamate. The second feature is that the new amide bond is an *isopeptide* bond. The enzyme-controlled activation of the H₄-folyl-monoglutamate C₅ side chain carboxylate rather than the traditional C₂ carboxylate means that the amide formed on attack by the amine of free glutamate is a γ -carboxamide not an α -carboxamide linkage. The addition of one to six additional glutamates all proceed *via* the same regiochemistry to yield a chain of γ -linked oligoglutamates.

The isopeptide linkages in the folyl oligoglutamyl coenzyme forms are sufficient evidence to predict that those peptide bonds were synthesized nonribosomally. The comparable regiochemistry of the γ -Glu-Cys isopeptide bond in glutathione is likewise a measure that glutathione is also not made on ribosomes. Ribosomal factories only make *peptide not isopeptide* bonds (see ubiquitylations with their isopeptide linkages to N₆ of lysine residues as *posttranslational* modifications in Chapter 16). In contrast, we will observe that the nonribosomal peptide synthetase assembly lines, multienzyme complexes that are RNA-independent in building peptide sequences into products, do make isopeptide as well as peptide bonds and even oxoester bonds.

15.4 Peptide Bond Initiation, Elongation and Termination at the Ribosomal RNA Machine

The many facets and intricacies of protein biosynthesis, mediated by the nanomachinery of ribosomes, studied in purified and functionally reconstituted forms by now for some 70 years, are far beyond the scope of our comments here. Introductions to the subject are available in standard biochemistry textbooks.

We summarize only peptide bond-forming events and strategies, without delving

into the issues of particular messenger RNA structures, structure and function of modified bases in tRNAs, efficiency of anticodon–codon trinucleotide base-pair recognition, rates of peptide bond formation, cotranslational folding and myriad other important aspects of protein biosynthetic machinery.

Instead, we limit our analysis to peptide bond formation (nitrogen) chemistry during initiation, elongation and termination events (schematized in [Figure 15.5A](#)).

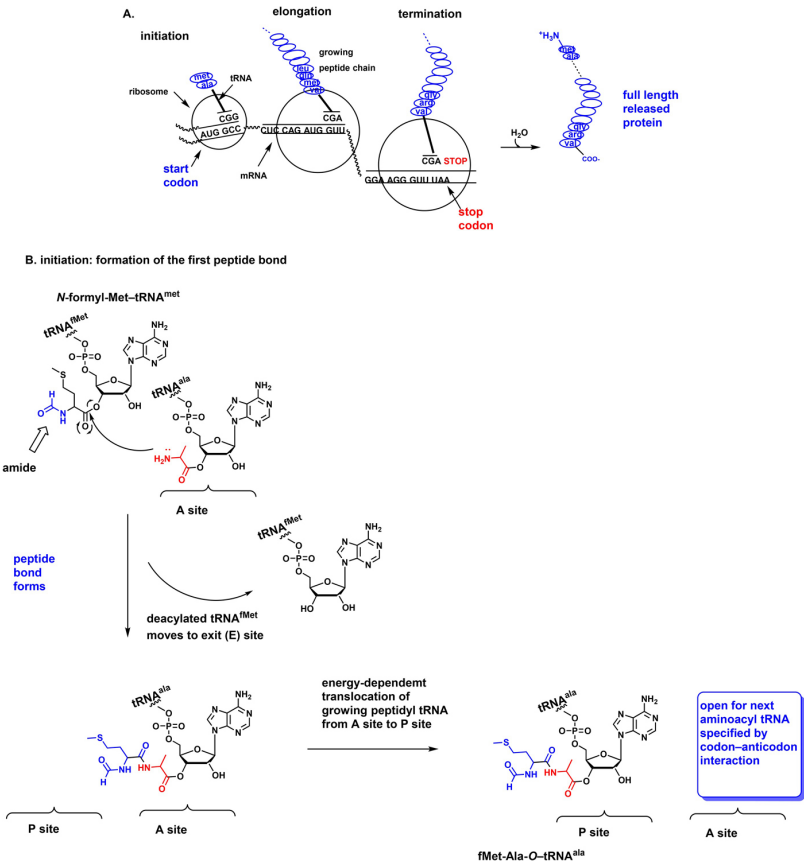


Figure 15.5 (A) A schematic overview of peptide chain initiation, elongation and termination stages at the ribosome. (B) The initiator fMet–tRNA^{fmet} is chaperoned into the P site of the peptidyl transferase center in the ribosomal large subunit. Initiation: formation of the first peptide bond and movement of the dipeptidyl-tRNA^{ala} to the P site.

15.4.1 Ribosomal Initiation of Peptide Chain Formation

Among the many aspects of protein chain initiation at the peptidyl transferase center of the bacterial ribosome, including setting and checking the codon–anticodon triplet base pairing match, conditional GTPase activity of the protein chaperone and activating the necessary conformations for activity, are the differentiation of the donor and acceptor aminoacyl tRNAs. (Eukaryotic ribosomes follow equivalent logic but with more initiation factors and controls on ribosomal subunit complex formation.)

Both aminoacyl groups are in oxoester linkage to their cognate tRNA, placed in

register by the two codon (mRNA)–anticodon (tRNA) pairs. The first codon in bacterial protein biosynthesis is always an AUG, calling for an initiator methionyl-tRNA. That methionyl residue is always *N*-formylated, thus cloaking the amino group's nucleophilicity. The initiator fMet-tRNA^{fMet} is guided to the peptidyl site in the active peptidyl transferase domain of the large ribosomal subunit (Figure 15.5B). That *N*-formylated methionyl oxoester can only act as the electrophilic aminoacyl acceptor via its oxoester at the carboxyl group (Figure 15.5B). In many proteins an alanine is the second residue, so an alanyl-tRNA^{ala} can act either as nucleophile at the amino group (when deprotonated) or as electrophile at the carboxyl-*O*-tRNA group. To pair with the *N*-formyl-methionyl-tRNA^{fMet} for peptide bond formation, the alanyl moiety can only act productively as a nucleophile (Figure 15.5B). The condensation of an oxoester with an amine (fMet with Ala) to the peptide (amide) bond is downhill energetically (Figure 15.5B) and proceeds to completion.

15.4.2 Peptidyl Chain Elongation at the Ribosome

All subsequent peptide bond-forming steps involve the growing peptidyl chain as an acyl oxoester electrophilic donor to the nucleophilic amine of an incoming aminoacyl-*O*-tRNA monomer. Thus, directionality of chain growth is ensured since only the aminoacyl monomers contain a nucleophilic nitrogen. Protein biosynthesis proceeds from the protein N-terminus to the C-terminus unidirectionally because of that chemical underpinning. Figure 15.6 illustrates the peptide bond-forming step for ribosomal condensation of a dipeptidyl fMet-Ala-tRNA^{ala} with a Ser-tRNA^{ser} as a chain elongation step that moves the newly generated tripeptidyl unit to the tRNA^{ser} as fMet-Gly-Ser-tRNA^{ser}.

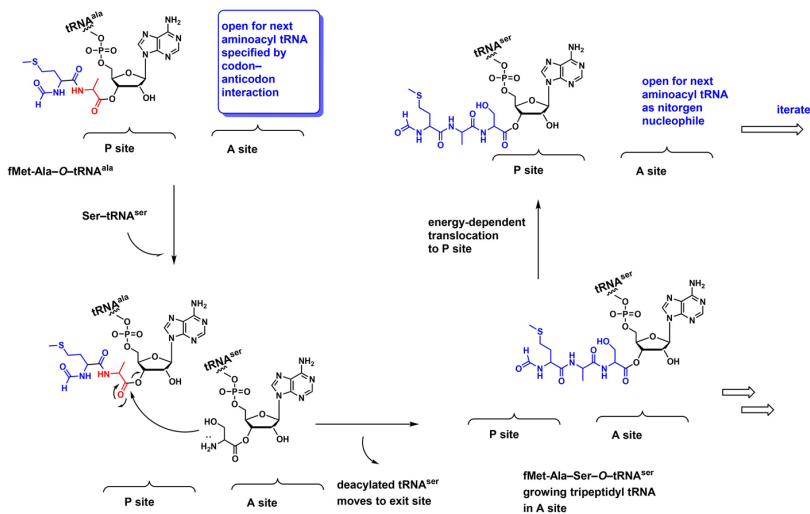


Figure 15.6 Peptide chain elongation; formation of the second peptide bond (chain elongation from dipeptidyl tRNA^{ala} to tripeptidyl tRNA^{ser}) followed by energy dependent shuttling to the adjacent P site.

Attendant to each chain elongation peptide-bond forming step, and actually preceding it, is the GTP-dependent movement of the growing chain between the A (aminoacyl) and P (peptidyl) sites. In any condensation step, including the first one, the peptidyl chain-*O*-tRNA resides in the P site. The incoming monomeric aminoacyl-

tRNA with its free amino group binds to the A site. That is the productive orientation of peptide bond formation. The growing peptidyl chain is translocated to the tRNA that was in the A site by GTP hydrolysis.

Before the next peptide bond can be formed, the nascent peptidyl-*O*-tRNA chain must be moved over to the P site, mechanical movement driven by the energy released as GTP is hydrolyzed. This frees up the A site for the next aminoacyl-*O*- to enter and then react. The reciprocating shuttle of the growing peptidyl chain from the P site to the A site in each peptide-bond forming step, and then back from the A site to the P site in time for the next condensation step is energy-dependent. That process of iterative shuffling goes on until the full-length chain has been reached. For a one hundred residue protein there are 99 such iterative ratcheting of the peptidyl chains between the A and P sites (Figures 15.5B and 15.6) (and 99 GTPs hydrolyzed). The rate of ribosomal peptide chain extension is about ten peptide bonds per second, so the iterative site switching is complete within milliseconds.

The largest human protein, titin has about 35 000 amino acids, weighing in at approximately 3.9 million Daltons per chain. Fungi make cyclosporin synthase, a hybrid polyketide-nonribosomal peptide synthase (see Section 15.5) with 11 modules in one protein of about 15 000 amino acids long. ¹² Among the smallest ribosomally synthesized proteins is the bacterial microcin C, a seven-residue peptide encoded by a gene only 21 nucleotides long. ¹³ Thus, the number of peptide bonds in proteins made on ribosomes in response to mRNAs can vary by a factor of a thousand.

15.4.3 Peptidyl Chain Termination at the Ribosome

Chain termination, likewise, is a highly coordinated event. Premature termination needs to be suppressed to avoid accumulation of incomplete, nonfunctional and probably unfolded protein chains. In turn, when an mRNA message has been read to the point where termination codons (Figure 15.5A) are reached, often as repeat stop codons, the growing peptide chain stalls. Termination factor proteins may be recruited to the ribosome to assist in a regiospecific disconnection of the full-length peptide chain from the last tRNA. The chemistry is *hydrolysis* of a peptidyl-*O*-3'-tRNA ester linkage (Figure 15.7). The release is always of a C-terminal *carboxylate*. (There are some peptide hormones, derived from larger proteins that terminate in carboxamides, but those come from specialized cleavage of a protein-glycine C-terminus *via* C₂-hydroxylation of the glycine moiety and unraveling of the resultant carbinolamine). The hydrolysis of an acyl(peptidyl)-*O*-ester to the free carboxylate is favored energetically (Figure 15.2B) so the hydrolytic release of full-length protein as the C-terminal carboxylate proceeds to completion, freeing up the ribosome for the next protein to be initiated, elongated and terminated.

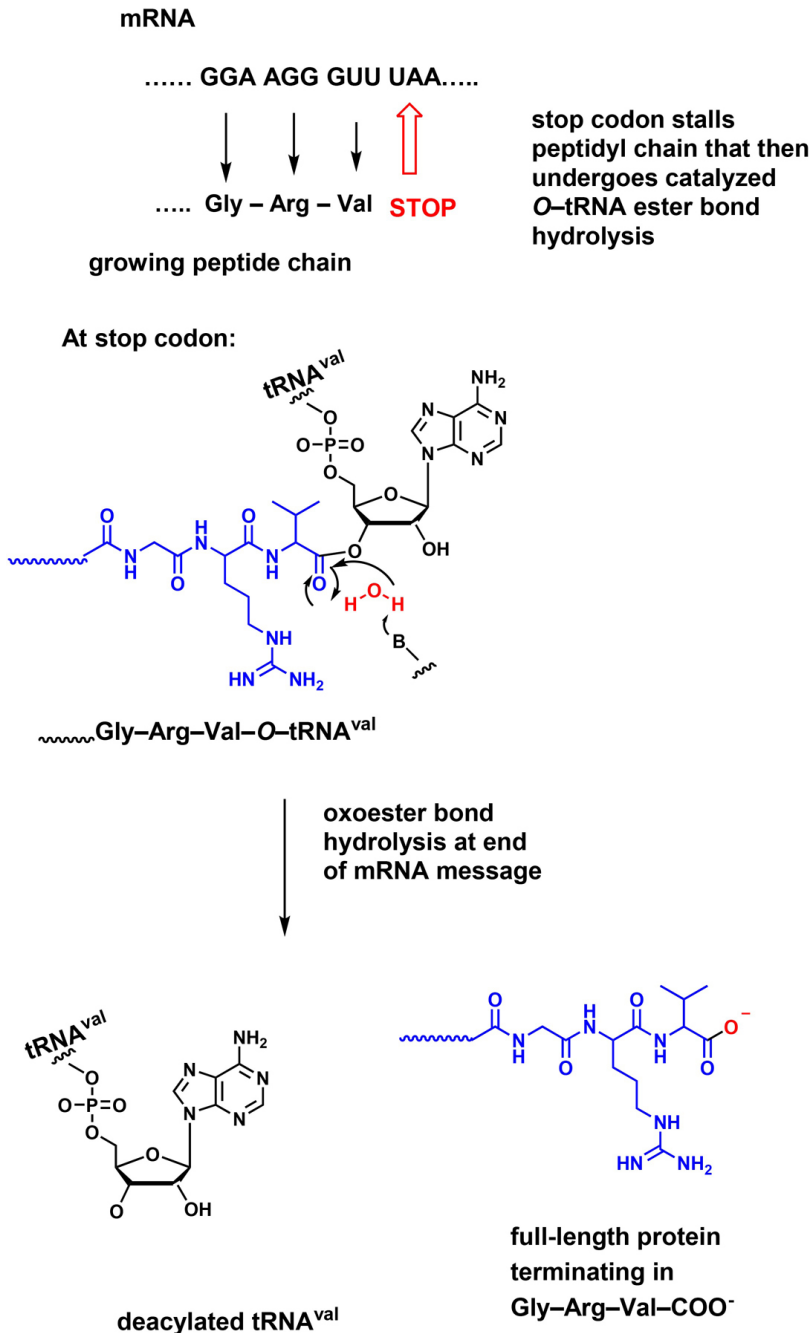


Figure 15.7 Termination: peptidyl chain terminates when the mRNA message shows a stop codon (here UAA). The peptidyl chain stalls at that point and undergoes catalyzed hydrolysis of the oxoester linkage with release of the deacylated tRNA and the full-length protein terminating in a carboxylate anion.

The example shown in [Figure 15.7](#), following the cartoon of [Figure 15.5A](#) reaches

the end of the mRNA-encoding message as the UAA stop codon appears at the end of the mRNA coding sequence. The growing protein chain, with the last three amino acids Gly–Arg–Val is still covalently attached as an oxoester to the 3'-OH of the terminal AMP moiety of the tRNA^{val}. Hydrolysis is catalyzed to cleave the oxo ester and release the protein with a C-terminal free carboxylate group (in this case the last three amino acids are –Gly–Arg–Val–COO⁻).

15.5 Chemical Logic for Nonribosomal Peptide Synthetase Assembly Lines

In addition to the small peptides made by soluble ATP grasp enzyme family members (Figure 15.2) where ATP is cleaved to ADP and the aminoacyl phosphate intermediates, there is a much larger class of microbial nonribosomal peptides made by the ATP to AMP + PPi cleavage mode. 14:15 This is the same ATP cleavage pattern practiced by the aminoacyl tRNA synthetases for ribosomal machinery. Indeed, the nonribosomal peptide synthetases (NRPS) have similar adenylation domains or subunits that likewise make aminoacyl-AMPs as tightly bound intermediates.

From that common point of the aminoacyl-AMP mixed aminoacyl-phosphoric anhydrides, ribosomal and nonribosomal nanomachinery and logic diverge (Figure 15.8). Whereas cognate tRNAs are substrates for the second half reactions of the amino acyl tRNA synthetases, as described in Section 15.4, in net aminoacyl transfers to oxygen nucleophiles to form amino acyl-oxoester bonds to tRNAs, the NRPS assembly lines transfer the amino acyl groups to thiolate nucleophiles, forming aminoacyl thioesters instead.

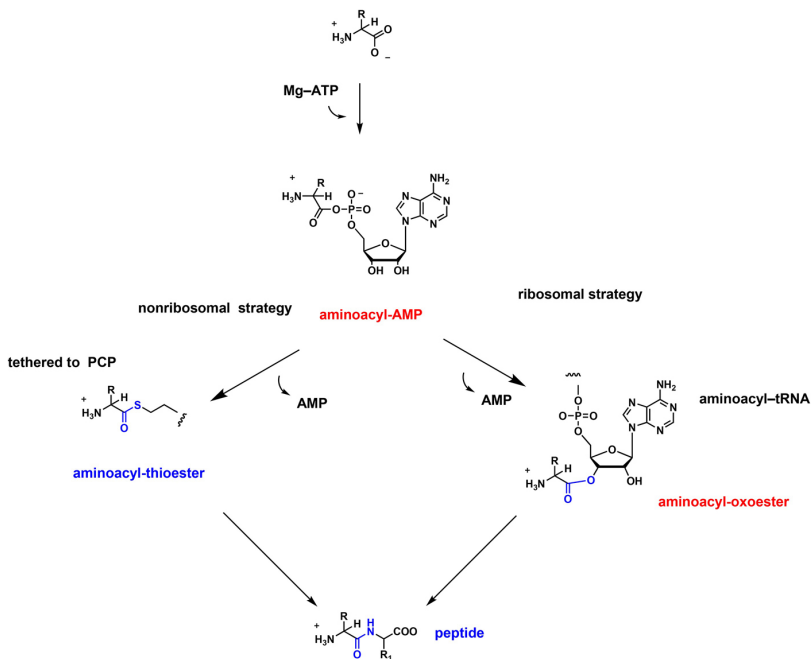


Figure 15.8 Both ribosomal and nonribosomal peptide biosynthesis starts with the common logic of activating each amino acid carboxylate as the aminoacyl-AMP

mixed anhydrides, then strategies diverge. NRPS machinery covalently tethers the amino acids and growing peptidyl chain as a series of peptidyl or aminoacyl-S-PCPs. Ribosomal machinery involves chaperoning the aminoacyl tRNA oxoesters to the ribosomal peptidyltransferase center. Both strategies result in peptide bond formation.

Furthermore, the thiol moieties are not free thiol small molecule substrates. The thiolate is part of a phosphopantetheinyl arm (Ppant) (half of coenzyme A: [Figure 15.9](#)) covalently tethered to a serine residue in an 8–10 kDa domain known as a peptidyl carrier protein (PCP). [16](#) Thus, the aminoacyl group becomes covalently tethered as a thioester to the Ppant arm of a carrier protein. Sometimes, *e.g.* in [Figure 15.10](#), the three letter PCP acronym is replaced by a one letter T (for thiolation domain).

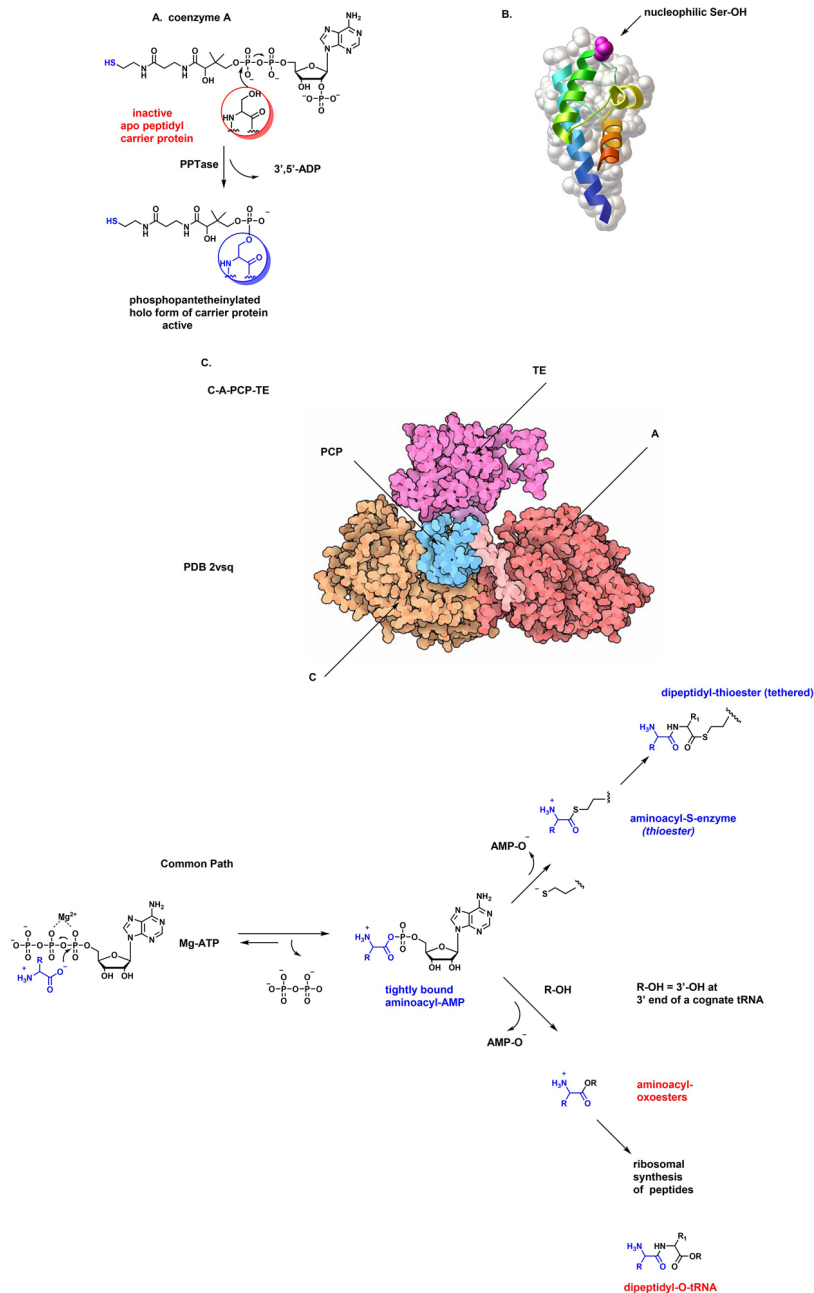


Figure 15.9 Peptidyl carrier proteins in NRPS assembly lines. (A) Enzymatic conversion of inactive apo form of peptidyl carrier protein (PCP) to active phosphopantetheinylated holo form by action of the posttranslational enzyme phosphopantetheinyl transferase using coenzyme A as substrate. (B) Peptidyl carrier proteins are small 80–100 amino acid residue proteins with a four-helix bundle core and a specific serine–OH side chain that acts as a nucleophile in conversion of the apo to the holo form (PDB 101, 2014). (C) The seventh and

terminal module of the *Bacillus* NRPS for the heptapeptide surfactin. This termination module has four domains, C–A–PCP–TE, as shown in the figure (PDB 101, 2014). Also shown is the partition of the common aminoacyl AMP between the ribosomal path to a dipeptidyl-*O*-tRNA vs. the nonribosomal path to a dipeptidyl thioester enzyme adduct.

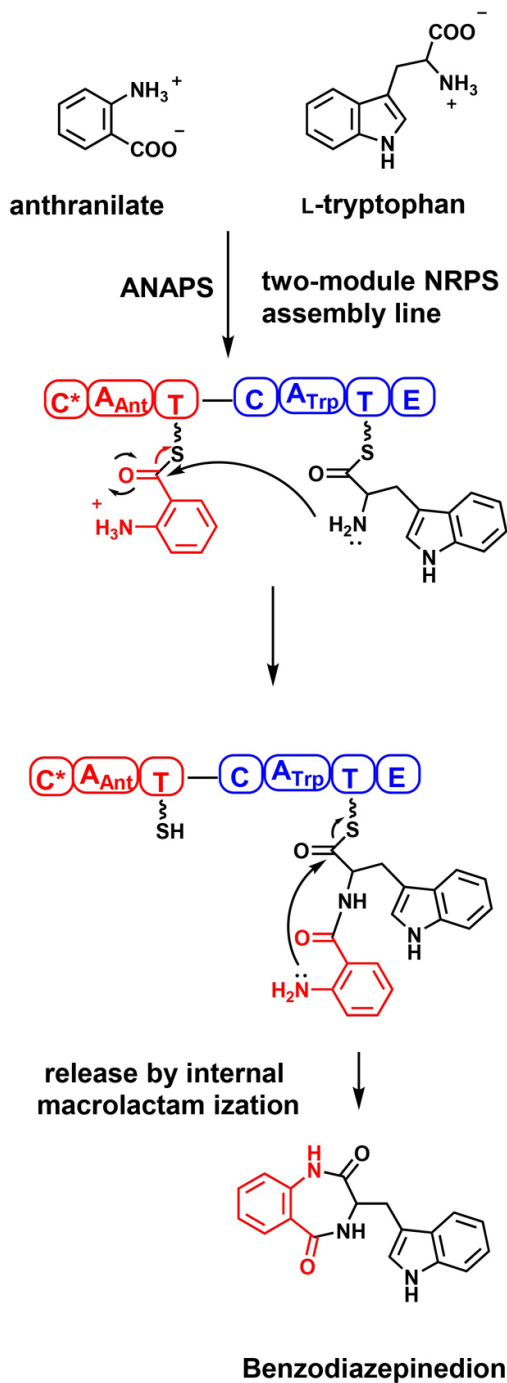


Figure 15.10 Schematic of a two-module fungal NRPS assembly line to make *R*-benzodiazepinedione from anthranilate and tryptophan. C* is an inactive condensation domain, E is an epimerase domain. T indicates a thiolation domain (the same as the PCP domain but with a one-letter code). Chain release is by

internal capture of the dipeptidyl thioester by the anthranilate residue free amino group to generate the unusual 6,7 fused heterocycle (a ring-expanded diketopiperazine).

The growing chain will involve a series of those aminoacyl-S-enzymes, with each amino acid specified by the recognition specificity of the adenylation domains. The initial peptide bond formation is then pairing of two aminoacyl-S-enzymes to form the dipeptidyl-S-enzyme with translocation of the growing peptidyl chain to the downstream PCP-S-Ppant arm.

In the specific example shown in [Figure 15.10](#) a fungal two-module assembly line builds the fused 6,7-bicyclic benzodiazepinone. ¹⁷ The A_{ant} adenylation domains selects anthranilate, activates it as anthranilyl-AMP and then covalently tethers it as the anthranilyl-S-pantheinyl thioester arm of the first T domain in the assembly line. In the second module A_{trp} accordingly binds, activates and installs L-Trp as the aminoacyl-S-T thioester. The peptide bond-forming condensation is catalyzed by the C domain of module two (C* in module one is inactivated by mutation) with the upstream anthranilyl carbonyl as an electrophile and the downstream Trp-NH₂ as nucleophile. The resultant Ant-Trp- dipeptidyl thioester is on the T₂ domain.

Chain release, a termination step, is unusually direct in this diazepinedione-forming system. Release occurs *via* a diketopiperazine-like (see Section 9.5.4) internal capture of the activated Trp carbonyl by the free NH₂ of the anthranilyl moiety. Because anthranilate is a β-amino acid rather than a conventional α-amino acid, the product ring system is an unusual 6,7-fused benzene-diazepinedione. The use of anthranilate, a nonproteinogenic monomer signifies that the product was not made on ribosomes.

The amino acid sequence of the growing peptide chain on an NRPS assembly line is *not* determined by information directly coded in DNA and read out as messenger RNA. Instead, the sequence of amino acids added is determined only by the specificity of the adenylation (A) domains and their amide-forming partner condensation (C) domains.

The independence from ribosomal strictures and the freedom from the specificity imposed by the aminoacyl tRNA synthetases for the 20 proteinogenic amino acids also means escape from the canonical 20 amino acid building blocks. Hallmarks of peptides from NRPS biosynthetic machinery are nonproteinogenic amino acid residues embedded in peptide chains. ¹⁸ [Figure 15.11](#) depicts three such amino acids where the amino acid monomer scaffold is rearranged. In the three examples the amino group is not at C₂ but at C₃ (beta alanine from Asp decarboxylase action), C₄ (GABA from glutamate decarboxylase action) or C₆ (aminoadipate). We will examine the regiospecific activation of the C₆ carboxylate of aminoadipate in the next section on aminoadipyl-cysteinyl-D-valine synthetase. When each of these unusual amino acids are converted to the tethered aminoacyl-S-PCP and embedded in the elongating peptide chain they generate *isopeptide bonds* at the point where they have been inserted, a telltale sign of a nonribosomal peptide synthetase assembly logic.

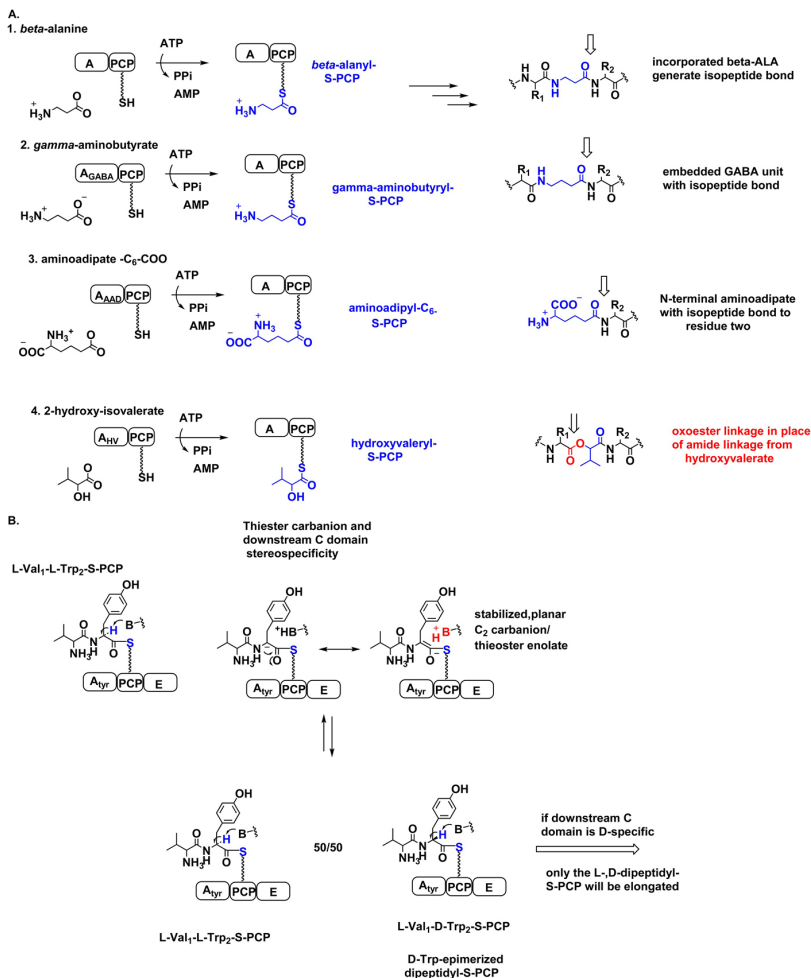


Figure 15.11 (A) Noncanonical building blocks for NRPS assembly lines include *beta*-alanine, *gamma*-aminobutyrate, aminoadipate (at the C₆ carboxylate) and the *hydroxy acid* 2-hydroxy-isovalerate. (B) Epimerization (E) domains catalyze “on assembly” line epimerization of peptidyl thioesters as shown. Only the D-configuration is recognized by the downstream C domain for peptide bond formation.

Figure 15.11 entry 4 also reveals that some NRPS adenylation (A) domains have evolved to recognize the neutral, planar 2-keto group of α -keto acids instead of the cationic NH_3^+ of amino acids. A paired NADH reductase domain (embedded in the protein assembly line) ensures that the activated keto acid is then chirally reduced to the L-hydroxyacyl-S-PCP thioester. 19,20 On incorporation into a growing NRP chain a novel oxoester bond replaces the normal peptide bonds, with different electronic and hydrolytic liability.

A second hallmark of NRPS assembly lines is that their peptide products often contain D-amino acid monomers, not seen in ribosomal protein biosynthesis where the chiral recognition barriers allow for L-amino acid building blocks only. Free D-amino acid metabolites are often rare and in limiting supply so most NRPS assembly lines, activate the more common L-amino acid and then epimerize C₂ to a d,l-mixture at the

stage of the aminoacyl-S-PCP or the corresponding peptidyl-S-PCP thioesters. The widely used antibiotic vancomycin contains three D-residues in its heptapeptide backbone: at residues 2, 4 and 5. [Figure 15.11B](#) suggests a mechanism for epimerization of Trp₂. The L-Val₁-L-Trp₂ tethered thioester is epimerized via a stabilized, delocalized C₂ carbanion with resonance contribution from the thioester enolate structure. Stereorandom return of a proton to C₂ should generate an equal mixture of starting L-L- and epimerized L,D-Val₁-Trp₂-S-PCP₂ diastereomers. If the downstream C domain in module 3 of the vancomycin synthetase is selective for a D-donor and an L-acceptor, only the L,D-dipeptidyl-S-PCP donor will be carried through to an L,D,L-tripeptidyl thioester as a growing chain intermediate.

15.5.1 NRPS Assembly Line Architecture

The individual catalytic domains of NRPS proteins (A = adenylation, C = condensation, PCP = peptidyl carrier protein, E = epimerase and TE = thioesterase) are typically strung together into functional modules that are autonomously folding and functioning as 110 to 160 to 200 kDa units, depending on domain composition. [21](#) A simplified schematic for the trimodular aminoadipyl-cysteinyl-D-valine (ACV) synthetase in penicillin biosynthesis, discussed in the next section is depicted in [Figure 15.12A](#). The individual domain and module boundaries can be computed from the encoding DNA sequences in bacterial genomes.

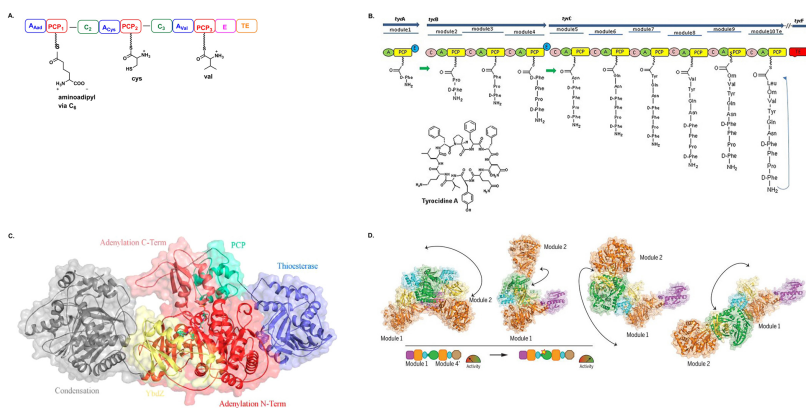


Figure 15.12A Two examples of microbial NRPS assembly lines. (A). The aminoadipyl-cysteinyl-D-valine (ACV) synthetase is a single protein of 400 kDa with 10 protein domains as indicated. ACV is the immediate precursor of penicillin type antibiotics. (B) The bacterial cyclic decapeptidyl lactam, tyrocidine A is made on a 3-protein 10-module, 32-domain assembly line. The most downstream domain the TE domain, catalyzes chain release by addition of the amino group of D-Phe₁ onto the activated carbonyl of Leu₁₀. (C) X-ray structure of the NRPS module EntF involved in assembly of the *Escherichia coli* siderophore enterobactin. (D) The dynamics of a two-module construct stitched together from modules 1 and 4 of the five-module gramicidin S synthetase. (B) Reproduced from [ref. 29](#), <https://doi.org/10.3389/fphar.2017.00828>, under the terms of the CC BY 4.0 license, <http://creativecommons.org/licenses/by/4.0/>. (D) Adapted from [ref. 30](#) with permission from PNAS, Copyright 2000 National Academy of Sciences, U.S.A. (C) Reproduced from [ref. 28](#) with permission from The American Society for Biochemistry and Molecular Biology, Copyright 2016. (D) Reproduced from [ref. 27](#) with permission from The American Association for the Advancement of Science, Copyright The Authors.

A typical assembly line for the nonribosomal decapeptide antibiotic tyrocidine is shown in [Figure 15.12B](#) where domain composition in each module allows prediction of what amino acids are activated by each A domain, when there is an epimerase E domain, and the presence and location of the most downstream TE (thioesterase) domain for decapeptide release. The genome sequence in this region predicts the presence of ten PCP domains one for each of the ten aminoacyl-S-PCP tethered, activated monomers for each peptide bond formation. ²² Three proteins, TycA, B and C, comprise the *tyc* assembly line, with one, three and six modules distributed in TycA, TycB and TycC, respectively.

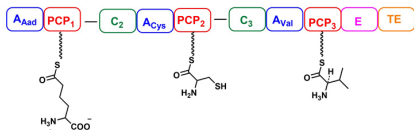
[Figure 15.12C](#) shows the X-ray structure of the NRPS module EntF involved in assembly of the *Escherichia coli* siderophore enterobactin. It is a four domain (C-A-PCP-TE) module of 140 000 Dalton molecular weight. ²³ The relative orientation of the four domains is shown. Finally, [Figure 15.12D](#) depicts some of the dynamics of a two-module construct stitched together from modules 1 and 4 of the five-module gramicidin S synthetase.

As in analysis of the ribosome as a protein synthesis nanofactory, it is helpful to also analyze the underlying logic of NRPS assembly lines for chain initiation, chain elongation, and chain termination. There are different domain compositions in initiation vs. elongation vs. chain termination modules, consistent with different architectural and chemical roles during peptide chain growth. Analysis of the paradigmatic three-module fungal protein ACV synthetase from fungal penicillin producers illustrates the core principles of NRPS logic ([Figure 15.13](#)).

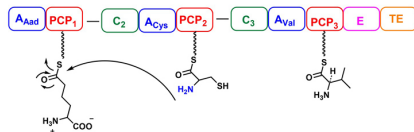
1. trimodular ACV synthetase: all three PCPs in holo phosphopantetheinyl form



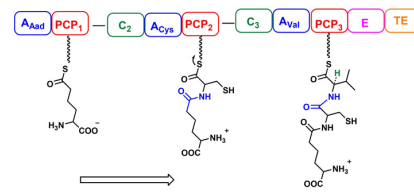
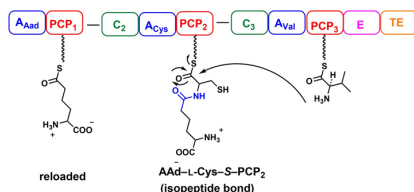
2. covalent loading of aminoadipate (AAd), L-cysteine, L-valine



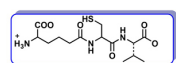
3. first peptide bond catalyzed by C₂ domain



4. second peptide bond catalyzed by C₃: peptide chain migrates as amide bond forms free PCP domains reload

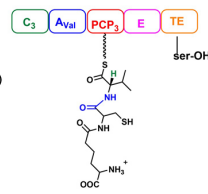


5. epimerization (see Figure 15.11B) for mechanism



aminoadipyl-L-Cys-D-Val = ACV

6. transfer of L-L-D-tripeptidyl chain to TE active site Ser-OH, then hydrolysis



B.

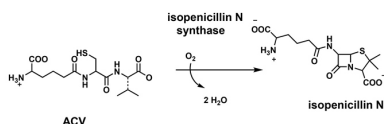


Figure 15.13 ACV synthetase: three modules and ten domains in one protein assembly line. (1) priming of apo to holo forms on all three PCP domains; (2) selection, activation with ATP and then covalent loading of the three amino acids on the respective PCP domains; (3) peptide bond #1 catalyzed by C₂ domain; (4) peptide bond #2 catalyzed by C₃ domain; (5) epimerization by E domain (see Figure 15.11B for proposed mechanism); (6) transfer of L-L-D-tripeptidyl chain from PCP₃ to Ser-OH of TE domain, preparatory to hydrolytic release. (B). The next enzyme in this fungal pathway is isopenicillin N synthase, creating the beta lactam strained chemical warhead of this antibiotic class.

15.6 ACV Synthetase: A Three Module NRPS Assembly Line

The enzyme ACV synthetase is the direct feeder enzyme of aminoadipyl-cysteinyl-D-valine to the isopenicillin N synthase, creating the active β -lactam antibiotic **24** (Figure 15.12). The presence of both aminoadipate, a nonproteinogenic amino acid, and its C₆ isopeptide bond already signaled that ACV was made outside of the ribosomal machinery.

After cloning, isolation, purification, structural and catalytic activity characterization, ACV synthetase turns out to have all its catalytic sites on a single 400 kDa protein chain (that is ribosomally produced). There are three modules, containing a total of 10 semi-autonomously folding domains as depicted in Figure 15.12. There is an initiation module, an elongation module and a termination module, containing two, three and five domains, respectively. We will examine each in turn for composition and function.

There is also the need to prime each of the three embedded PCP domains, one in each module, to convert the inactive apo forms of the carrier protein domains to the posttranslationally phosphopantetheinylated forms to provide the thiolate nucleophiles on which the amino acids and growing peptide become covalently tethered **16** (Figure 15.13).

15.6.1 ACV Synthetase Initiation Module

Once all three PCP domains have been matured posttranslationally to the Ppant forms (Figure 15.13), a catalytic cycle in an NRPS assembly line can begin. Module 1 has two domains, a 50 kDa adenylation domain and a 10 kDa PCP domain. This adenylation domain is specific for recognizing the unusual amino acid metabolite aminoadipate. Furthermore, it catalyzes attack of the C₆-carboxylate, not the C₂-carboxylate, on P _{α} of Mg-ATP to produce tightly bound aminoadipyl-C₆-AMP. That activated aminoadipyl C₆ carbonyl is then captured by the Ppant-S⁻ arm and thereby transferred to the PCP₁ domain as the aminoadipyl-C₆-S-PCP₁ thioester (Figure 15.13A).

15.6.2 ACV Synthetase Elongation Module

In the three-module ACV assembly line there is only one elongation module: module 2. This contains three domains: the condensation₂, adenylation₂ and PCP₂ domain. (There is no condensation domain in the initiation module.) The A₂ domain is specific for binding L-cysteine, activating it as the cysteinyl-AMP adduct and transferring the now activated cysteinyl moiety to the thiol end of the Ppant domain of PCP₂. At this juncture, there are two aminoacyl-S-PCP thioesters, upstream aminoadipyl-S-PCP₁ and downstream L-cysteinyl-S-PCP₂. The C₂ condensation domain then acts as a peptide synthase, bringing the two PCP domains in apposition in its active site where the deprotonated amine of the Cys-S-PCP₂ acts as a nucleophile on the aminoadipyl C₆-S-PCP₁. The result is transfer of the aminoadipyl group to PCP₂ attendant with (iso)peptide bond formation to give the dipeptidyl aminoadipyl-(C₆)-L-Cys-S-PCP₂ (step 3 of Figure 15.13A). Chain growth on NRPS assembly lines is unidirectional: from upstream to downstream PCP domains.

15.6.3 ACV Synthetase Termination Module

There are five domains in module 3 of ACV synthetase. This is the chain termination module. Three of the five domains C₃, A₃ and PCP₃ mirror the three cognate domains of module 2. Two are additional: the epimerase (E) domain and the most downstream of all 10 domains, the thioesterase (TE) domain. Module 3 catalyzes the second peptide bond formation (step 4), epimerization of the valyl moiety (step 5) and hydrolysis by the TE domain (step 6).

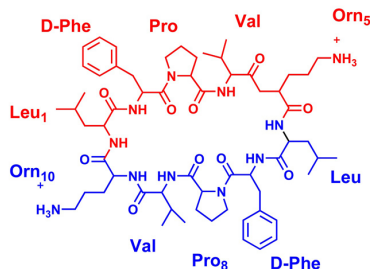
The A₃ domain is specific for activating L-valine, even though D-valine is residue 3 in the released ACV tripeptide. That A domain makes L-valyl-AMP and harvests the thermodynamic activation with improved kinetic stability as L-Val-S-PCP₃. There is some uncertainty about the timing of the C₃ vs. the E₃ domain. Evidence in other NRPS assembly lines indicates that C domains act before E domains. If so, then C₃ would catalyze the attack of the deprotonated amino group of L-Val-S-PCP₃ on the activated thioester carbonyl of AAd-Cys-S-PCP₂ (step 4 of [Figure 15.13A](#)). The product will be AAd-L-Cys-L-Val-S-PCP₃.

The tripeptidyl scaffold has been formed and is installed on module 3 next to the TE domain. However, the chirality of the valine residue is still L-, not D-. Now, the E-domain acts, abstracting the C₂-H of the L-valyl moiety as a proton. The resulting carbanion is stabilized, and thus accessible, as the thioester enolate (see [Figure 15.11B](#) for mechanism of “on assembly line” epimerization). Stereorandom return of the proton would give a 1:1 mix of L-valyl and D-valyl moieties. If the TE domain is specific for the conformer of the AAd-L-Cys-D-Val tripeptide then flux to the correct product stereochemistry will ensue.

The tenth domain in the ACV synthetase and the most downstream domain whenever it is present in an NRPS assembly line is the TE domain. It acts to transfer the AAd-L-Cys-D-Val tripeptidyl chain from PCP₃ to the side chain serine-OH in the TE active site (step 6 of [Figure 15.13A](#)). At that point the TE domain acts as a hydrolase, releasing AAd-L-Cys-D-Val, the immediate substrate for the next enzyme isopenicillin N synthetase, a mononuclear nonheme iron radical catalyst ([Figure 15.13B](#)).

15.7 Nonribosomal Peptide Variants

There are many variants of NRPS assembly lines, with conditional domains, including N-methylases, and variant TE domains that cyclize the full-length peptidyl chains into *cyclic macrolactams*, as in gramicidin S ([Figure 15.14](#)). As it happens, the NRPS assembly line for gramicidin S contains only five modules. ²⁵ The cyclic decamer is an exact repeat of two sets of five residues, indicating that the TE domain couples two identical tethered pentapeptide units, one on PCP₅ as a thioester, one on TE as an oxoester.

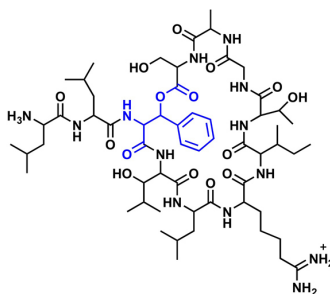


Gramicidin S

A five module assembly line
for a cyclic decapeptide
macrolactam

Head to Tail Dimerization
of Identical Pentapeptides

Lysobactin: undeca-peptidyl macrolactone



3-OH-Phe as building
block for **macrolactone**

lactone between residue 3-11

(D-Leu₁-Leu₂-OH-Phe₃-OH-Leu₄-Leu₅-D-Arg₆-
Ile₇-alloThr₈-Gly₉-H-Asn₁₀-Ser₁₁)

Figure 15.14 Two different types of cyclic peptides from NRPS assembly lines. Gramicidin S is a cyclic *decapeptide lactam* but the assembly line has only five modules. Inspection reveals that residues 1–5 and 6–10 are identical (Leu–D-Phe–Pro–Val–Orn), indicating internal dimerization of two pentapeptidyl intermediates, one on PCP₅ and the other on the TE in module 5. The second cyclic peptide is a *macrolactone* (cyclic ester not cyclic peptide) arising from capture of the activated Ser₁₁ carbonyl in the tethered linear undeca-peptidyl enzyme adduct by the 3-OH of the noncanonical building block 3–OH–Phe.

Some A domains of NRPS assembly lines have evolved to activate keto acids and reduce them to 2-hydroxy acids in place of amino acids to lead, for example, to peptide macrolactones. ¹⁹ Alternatively, macrolactones can form when β -hydroxyamino acids serve as NRPS monomers as in the case of lysobactin where a 3-OH-phenylalanine is the oxygen nucleophile in macrolactone formation (Figure 15.14). That macrolactone encompasses residues 3–11 of the 11mer scaffold.

The longest nonribosomal peptides built by an NRPS assembly line are 22 or 25 amino acids in syringopeptin22 and syringopeptin25, cyclic lipo-depsipeptides with long lipid tails that act as membrane conductance dissipators. ²⁶ NRPS machinery is largely relegated to microbes. There is one *bona fide* single cluster known in *Caenorhabditis elegans*.

It is possible that the energetic cost to build long protein assembly lines became outweighed by the investment of ATP to build modules. A typical C_n, A_n, PCP_n elongation module is in the order of 110 kDa (about 1000 amino acid residues). For a ten-module assembly line, with one initiation and one termination module, the size of the tyrocidine synthase (Figure 15.12B) on a three-protein array of 1, 3 and 6 modules would correspond to an assembly line with a protein molecular weight of 1 million Daltons, about 10 000 amino acid residues (9999 peptide bonds). All this to make a decapeptide. That decapeptide would need to have virtuoso biological functions to justify the energy to build the RNA-independent NRPS protein assembly line. Also, each NRPS assembly line is a highly customized nanomachine, not amenable to ready throughput to other peptide products. One finds 10mers but very few 20mers. When a novel amino acid building block is incorporated that is outside the proteinogenic amino acids and has some remarkable property (such as the warhead in penicillin antibiotics), the cost may be worth it.

15.8 Overview

Cells are replete with peptide bonds, perhaps 10¹⁰ in steady state in a mammalian cell. The chemical reaction is among the most straightforward in organic nitrogen chemical biology. This is amide formation where both reactants are amino acids. One central issue in all biological routes to amide bonds from amines and carboxylates is the need to activate the carboxylate anion of one of the partners to convert it to an electrophilic group that can be captured by the nucleophilic amine of a partner reactant. Amide biosynthesis from amine and carboxylate components is uphill thermodynamically and requires energy input to drive the equilibrium towards amide accumulation.

A molecule of the paradigmatic cellular energy currency, ATP, is spent for every amide or peptide bond formed. Two variants of ATP phosphoric anhydride side chain cleavage are employed, differentiated by attack of substrate carboxylate on P_γ (phosphoryl group transfer) or P_α (adenylyl transfer) of Mg-ATP. The resultant aminoacyl phosphate and aminoacyl AMP species are both mixed acyl-phosphoric anhydrides. Thermodynamic activation has been achieved but kinetic liability to hydrolysis can be problematic due to off target dissipation of the activated aminoacyl group. The ATP to AMP and PPi route dominates the energy input mode for all classes of cellular biopolymers.

Transfer of the activated (electrophilic) aminoacyl moiety to either a thiolate nucleophile or an oxygen nucleophile leads to aminoacyl thioesters or aminoacyl oxoesters, respectively. The amino oxoester strategy (in the specific context of aminoacyl-O-tRNAs) has won out in ribosomal machinery for making peptide bonds. That strategy and mechanism accounts for more than 99.99% of the cell's peptide bonds, made at the RNA catalytic peptidyl transferase center of ribosomes.

The parallel aminoacyl thioester strategy is used by nonribosomal peptide synthetase assembly lines, largely in bacteria and fungi. The thioesters are more activated thermodynamically than oxoesters and also more kinetically labile to hydrolysis. They are kept covalently tethered to phosphopantetheinyl arms on carrier protein domains in NRPS assembly lines.

The second major question in peptide bond formation in cellular contexts is how amino acid sequence is determined. For ribosomal protein biosynthesis, by far the great bulk of peptide bonds, this is the final third of the central dogma that DNA makes RNA makes proteins. The permanent triplet code in genes in DNA is translated into the complementary triplet code in messenger RNAs and those are the working blueprints for all approximately 25 000 different proteins in human cells. This is canonical macromolecular bioinformation machinery.

In contrast, the nonribosomal peptide synthetases of microbes are independent of mRNA in the sense that sequence specificity of product oligopeptides is not directly determined by triplet codon logic. The specificity and order of adenylation domains in modular assembly lines determines the product sequence.

NRPS machinery is not robust for making large peptides or proteins. The energy cost for building the assembly lines is high. On the other hand, the ability to incorporate noncanonical amino acid monomers and even to make hybrid peptide–polyketide assembly lines, offers unique bioactive small molecule niches to these assembly line scaffolds.

References

1. C. T. Walsh, B. Tu and Y. Tang Eight kinetically stable but thermodynamically activated molecules that power cell metabolism, *Chem. Rev.*, 2018, **118**, 1460-1494.
2. R. Geige The early history of tRNA recognition by aminoacyl-tRNA synthetases, *J. Biosci.*, 2006, **31**, 477-488.
3. P. Berg and E. Offenbach An enzymatic mechanism for linking amino acids to RNA, *Proc. Natl. Acad. Sci. U. S. A.*, 1958, **44**, 78-86.
4. F. Crick The origin of the genetic code, *J. Mol. Biol.*, 1968, **38**, 367-379.
5. M. Ibba and D. Soll Aminoacyl-tRNA synthetases, *Annu. Rev. Biochem.*, 2000, **69**, 617-650.
6. M. Y. Galperin and E. Koonin A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity, *Protein Sci.*, 1997, **6**, 2639-2643.
7. M. V. Fawaz, M. Topper and S. M. Firestone The ATP-grasp enzymes, *Bioorg. Chem.*, 2011, **39**, 185-191.
8. H. Arulanantham, N. Kershaw, K. S. Hewitson, C. E. Hughes, J. E. Thirkettle and C. J. Scholfield ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid, *J. Biol. Chem.*, 2006, **281**, 279-287.
9. Y. Kitamura, A. Ebihara, Y. Agari and A. Shinkai Structure of D-alanine-D-alanine ligase from *Thermus thermophilus* HB8: cumulative conformational change and enzyme-ligand interactions, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2009, **65**, 1098-1106.
10. L. Chen, H. Qi, T. A. Garrow, Y.-J. Choi and B. Shane Purification and properties of human cytosolic folylpolyglutamate synthetase and organization, localization, and differential splicing of its gene, *J. Biol. Chem.*, 1996, **271**, 13077-13087.
11. C. T. Walsh and Y. Tang, *The Chemical Biology of Human Vitamins*, Royal Society of Chemistry, London, 2019, p. 446.
12. H. Kleinkauf, J. Dittman and A. Lawen Cell-free biosynthesis of cyclosporin A and analogue, *Biochim. Biophys. Acta*, 1991, **50**, S219-S224.
13. J. E. Gonzalez-Pasto, J. San Milan and F. Moreno The smallest known gene, *Nat. Rev. Microbiol.*, 194, **369**, 281.
14. M. A. Fischbach and C. T. Walsh Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms, *Chem. Rev.*, 2006, **106**, 3468-3496.
15. D. Schwarzer, R. Finking and M. A. Marahiel Nonribosomal peptides: from genes to products, *Nat. Prod. Rep.*, 2003, **30**, 275-287.
16. R. H. Lambalot, A. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla and C. T. Walsh A new enzyme superfamily—the phosphopantetheinyltransferases, *Chem. Biol.*, 1996, **3**, 923-936.
17. C. T. Walsh, S. Haynes, B. D. Ames, X. Gao and Y. Tang Short pathways to

- complexity generation: fungal peptide alkaloid multicyclic scaffolds from anthranilate building blocks, *ACS Chem. Biol.*, 2011, **8**, 1366-1382.
18. C. T. Walsh, R. O'Brien and C. Khosla Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds, *Achie*, 2018, **52**, 7098-7124.
 19. D. A. Alonzo, N. A. Magarvey and T. Scheming Characterization of cereulide synthetase' a toxin producing macromolecular machine, *PLoS One*, 2015, , e0128569.
 20. N. A. Magarvey, M. Ehling-Schultz and C. T. Walsh Characterization of the cereulide NRPS *alpha*-hydroxy acid specifying module" activation of *alpha*-keto acids and chiral reduction on the assembly line, *J. Am. Chem. Soc.*, 2006, **128**, 10698-10699.
 21. B. M. Miller and A. Gulick Structural biology of non-ribosomal peptide synthetases, *Methods Mol. Biol.*, 2017, **1401**, 3-29.
 22. H. D. Mootz and M. Marahiel The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains, *J. Bacteriol.*, 1997, **179**, 6843-6850.
 23. J. Reichert, M. Sakitani and C. T. Walsh Characterization of EntF as a serine-activating enzyme, *Protein Sci.*, 1992, **1**, 549-566.
 24. J. Martin ACV synthetases in *beta*-lactam producing microorganisms, *J. Antibiot.*, 2000, **53**, 1008-1021.
 25. T. Stachelhaus and M. Marahiel Modular structure of peptide synthetases revealed by the dissection of multifunctional enzyme GrsA, *J. Biol. Chem.*, 1995, **270**, 6163-6169.
 26. V. Vassilev, P. Lavermicocca, D. Di Giorgio and N. S. Iacobellis Production of syringomycins and syringopeptins by *Pseudomonas syringae* pv. *atrofaciens*, *Plant Pathol.*, 1996, **45**, 316-322.
 27. J. M. Reimer, M. Eivaskhani, I. Harb, A. Guarne, M. Weigt and T. M. Schmeing Structures of a dimodular nonribosomal peptide synthetase reveal conformational flexibility, *Publ. Am. Assoc. Adv. Sci.*, 2019, **366**(6466), eaaw4388.
 28. B. R. Miller, E. J. Drake, C. Shi, C. C. Aldrich and A. M. Gulick Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture, *J. Biol. Chem.*, 2016, **291**, 22559-22571.
 29. S. Agrawal, D. Acharya, A. Adholeya, C. J. Barrow and S. K. Deshmukh Nonribosomal Peptides from Marine Microbes and their Antimicrobial and Anticancer Potential, *Front. Pharmacol.*, 2017, **8**, 828.
 30. H. D. Mootz, D. Schwarzer and M. A. Marahiel Construction of hybrid peptide synthetases by module and domain fusions, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**(11), 5848-5853.

Nitrogen Side Chains in Enzyme Catalysis and Posttranslational Modifications

In a sense all of protein chemistry revolves around the precise structure and folding of the peptide backbone to position side chains in specific packing and orientations, stabilizing secondary, then tertiary and quaternary structural elements in proteomes. Even within the subclass of proteins with catalytic activity, the subject of how thousands of enzyme catalysts operate with both specificity and rate acceleration is beyond the scope of this nitrogen-centered approach. This caveat extends to the more than 200 types of stable covalent posttranslational modifications (PTMs) on client proteins both on side chains and backbone connectivities. Treatises have been written over many years on both the origins of enzymatic catalysis [1](#) and the nature and mechanisms of posttranslational modification enzymes. [2](#)

16.1 Six Proteinogenic Amino Acid Building Blocks with Side Chain Nitrogen Atoms

In this second chapter in the section on nitrogen proteomes we summarize major observed reactivities of six of the twenty proteinogenic monomers that end up with nitrogen-containing side chains in folded proteins of all categories. In turn, those six residues are in two subcategories based on whether there is or is not a titratable pK_a value, equilibrating the ammonium cation form and a free base amine form of that nitrogen side chain group.

[Figure 16.1](#) shows that the imidazole ring of histidine, the N_ϵ -ammonium side chain of lysine and the guanidinium group of arginine have pK_a values of 6.0, 10.8 and 12.5, respectively as the free amino acids. (Those pK_a values can be perturbed in particular microenvirons of folded proteins.) With a pK_a of 6.0 for the imidazolium ion, losing a proton to the neutral imidazole ring, histidine side chains are $>90\%$ in the neutral imidazole form at physiological pH values (approximately pH 7.2). Given that abundance [3](#) of the deprotonated, neutral form of the nitrogen(s) in the imidazole ring, one anticipates that histidine residues will be most effective in two modes of reactivity ([Figure 16.2](#)), as basic groups to react with protons, and as nucleophiles to react with electrophiles beyond protons (phosphorus, carbon).

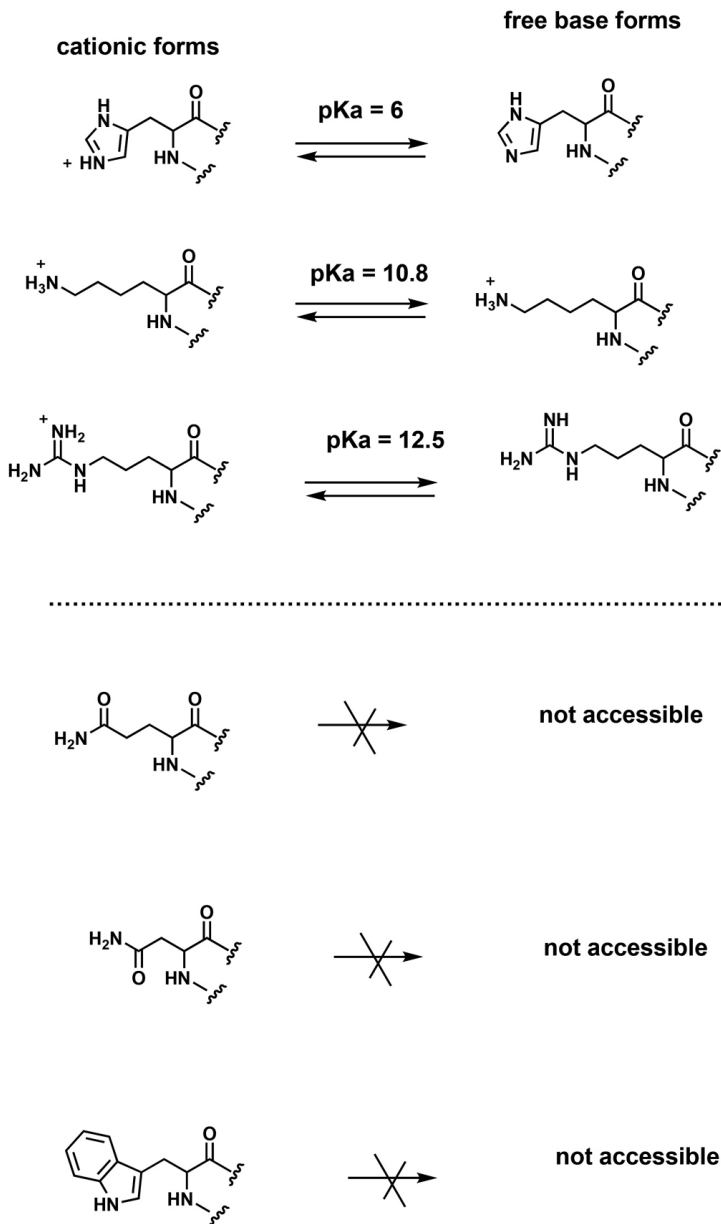


Figure 16.1 Six protein amino acid residues with side chain nitrogen functional groups: an imidazole, an amine, a guanidinium, two amides and an indole.

approximately 90% free base form at pH 7

1. imidazole nitrogen as a nucleophile

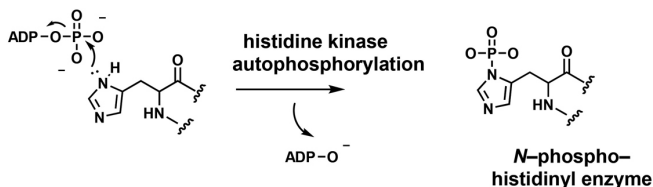
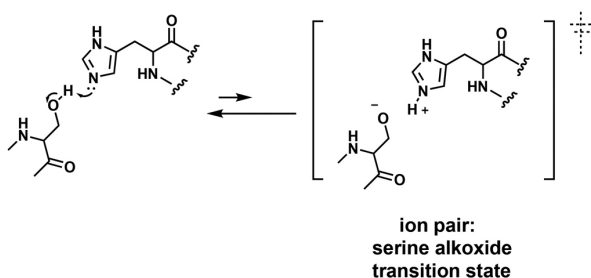
2. imidazole nitrogen as a base (towards H^+)

Figure 16.2 Deprotonated imidazole side chain nitrogen can act as a base or as a nucleophile.

The next closest pK_a to neutral pH is the $N_6-NH_3^+$ of lysine residues dissociating to a proton. The N_6-NH_2 where the lone pair on that nitrogen is available is a major side chain reactant, both for reprotonation as a base, and for capture of other electrophiles in reaction as a nucleophile. The third proteinogenic amino acid with a cationic nitrogen side chain at neutral pH is arginine. With a pK_a of 12.7 for the free amino acid, only one part in 10^5 is in the neutral guanidine form at neutral pH. That predominant bidentate guanidinium cation is crucial for electrostatic interaction with posttranslationally introduced phosphate groups to drive conformational changes. ³ Yet, we will observe that arginine has several modes of reactivity as a net nucleophile at the guanidinium moiety.

Thus, proteins contain three types of side chain nitrogens that can, in principle, act both as acids and bases and nucleophiles. These can be key functional groups in enzymatic catalysis. In complementary roles the nucleophilicity of the deprotonated forms are the basis for stable covalent posttranslational modifications as parts of protein substrates rather than the protein catalysts. The heterocyclic imidazole ring, the long chain C_6 amine of lysine, and the guanidinium group offer three highly differentiated nitrogen functional groups.

The remaining three nitrogen side chains (Figure 16.1) offer no titratable nitrogens within the range of biology, nor are the nitrogen atoms robustly nucleophilic. The asparagine and glutamine side chain carboxamides are a resonance-deactivated form of nitrogen. In particular, we have emphasized (Chapter 6) that the glutamine side chain carboxamide is utilized over and over as an unreactive carrier of latent NH_3 , releasable by hydrolytic glutaminase action at a target enzyme that uses nascent ammonia for

organic scaffold amination.

The next lower homolog of glutamine, asparagine, is not used as a ready source of latent ammonia equivalents. Remarkably though, nature has evolved a protein context, posttranslational *N*-glycosylation of thousands of eukaryotic proteins, where the carboxamido side chain nitrogen of Asn residues is coaxed into becoming a nucleophile, 4 as discussed in Section 16.6.

The sixth and final proteinogenic amino acid with a nitrogen-containing side chain is tryptophan with its indole ring. We spent the second half of Chapter 8 describing the innate reactivity of the N₁, C₂, C₃ enamine moiety of the pyrrole ring of the indole heterocycle in tryptophan and tryptamine as a potential or actual carbon nucleophile. That reactivity, the basis of thousands of plant and fungal indole alkaloid metabolites, is almost entirely suppressed in tryptophan residues in proteins. We will look at one example of indole ring C₂ mannosylation 5 in Section 16.8.

16.2 Transient vs. Stable Covalent Adducts on Protein Residue Nitrogen Side Chains

With this chapter's focus on reactivity of nitrogen side chain protein amino acid residues, we lump together catalysis with posttranslational modifications. This is notionally unusual, since enzyme catalytic cycles are typically complete in fractions of seconds while posttranslational modifications of proteins produce covalent side chain adducts that can last for hours to days in cells. Certainly, the role of the imidazole side chain of the histidine residue, acting as a *catalytic base*, 10–100 times per second in a protease active site, is not mimicked by single-turnover covalent adduct formation in a client protein away from the active site. However, the behavior of lysine, histidine and arginine, as *nitrogen nucleophiles* capturing electrophilic fragment of substrates, is the same whether occurring in milliseconds or 100 to 10 000 times slower. The lifetime of the covalent adducts is often the major distinction between a catalytic cycle of an enzyme and, for example, the posttranslational phosphorylation of a serine side chain in a protein kinase substrate.

This volume is neither a tome on enzymatic catalytic mechanisms nor a treatise on posttranslational modifications. Readers can access that literature elsewhere. This nitrogen-centered chemical biology focus is necessarily incomplete both in terms of active site side chains that participate in catalysis (serine–OH and cysteine–SH being the other major players) and the full scope of posttranslational modifications (again serine, threonine and tyrosine for phosphorylations and cysteine for up to seven types of reactions). Instead, this chapter stays focused on what kinds of basic and/or nucleophilic reactions the six nitrogen side chains of the proteogenic amino acid building blocks engage in that reflect their different nitrogen functional group reactivity.

16.3 Histidine as Acid–Base and Nucleophilic Catalyst

As stated above, the neutral form of imidazole side chains of histidine residues, given the pK_a value of 6.0, are the most abundant (in terms of mole fraction) of basic or nucleophilic nitrogen atoms that could participate in reactions with H⁺ as an electrophile (acting as a base) or with other electrophiles (acting as a nucleophile).

16.3.1 Base–Acid Catalysis in Serine and Cysteine Proteases

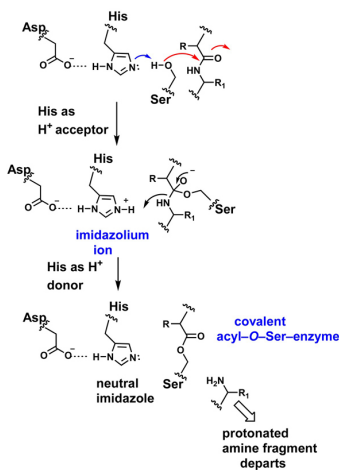
Perhaps the most extensively studied role of histidine as a general base catalyst occurs in the related families of “serine” and “cysteine” proteases. There are approximately 300 such proteases in human proteomes, expressing wide divergence in specificities for regioselectivity of peptide bonds hydrolyzed in client proteins.

In passing, we note that limited proteolysis, a key set of biological functions, constitutes one major category of posttranslational modification of protein substrates. These enzymes engage in covalent catalysis, using the active site “serine”–OH or “cysteine”–SH as catalytic nucleophiles (Figure 16.3A). The serine and cysteine proteases are two of four major mechanistic categories of proteases. The other two, “aspartyl” proteases and metalloproteases directly activate a cosubstrate water molecule to attack a client protein peptide bond and avoid covalent acyl enzyme intermediates.

A. first half reaction leads to acyl-O-Ser-enzyme covalent adduct

B. second half reaction is deacylation by water

1. active site histidine acts as base then acid, serine as nucleophile



2. active site histidine acts as base then acid, water molecule as nucleophile

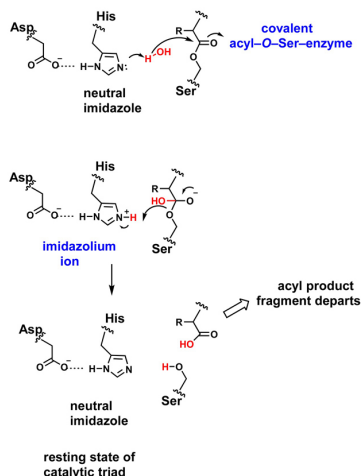


Figure 16.3 Serine proteases: active site histidine acts as a proton transfer catalyst (reciprocating base–acid catalysis) rather than as a nucleophile towards the amide carbonyl of substrates. Imidazole as base, then imidazolium ion as acid in both half reactions.

Paradigmatic members of both serine proteases and cysteine proteases **6** have a histidine imidazole side chain in the active site along with the serine–OH or cysteine–SH. Early on in historical analysis of mechanisms when it was clear that covalent acyl enzyme intermediates were part of these protease catalytic cycles, it was anticipated that the imidazole side chains would be the catalytic nucleophiles not the Ser–OH (pK_a approximately 14) or Cys–SH (pK_a approximately 8.5) base, on the basis of abundance of the nucleophilic forms.

That is definitively not the case for either Ser or Cys proteases. We examine the Ser protease strategy but it also holds for Cys proteases. The imidazole side chain acts as a base to initiate catalysis by proton transfer from the Ser–OH to generate the serine

alkoxide equivalent in the transition state (Figure 16.3). In the very next step, the resultant histidine imidazolium nitrogen now acts as a general acid catalyst, delivering a proton to the developing amine leaving group, avoiding discrete formation of the amine anion (pK_a approximately 35).

The acyl enzyme is the result, an oxoester, while the imidazole ring is back in the neutral state. The second half reaction begins when that imidazole acts as a base to a water molecule to create the hydroxide ion equivalent to act as nucleophile on the acyl enzyme carbonyl group. Immediately thereafter, the imidazolium ion then acts reciprocally as a general acid catalyst protonating the serine active site oxygen as it leaves as a neutral $-OH$ rather than the high-energy serine alkoxide.

The imidazole side chain in all these canonical serine and cysteine proteases acts reciprocally as a base, then an acid catalyst, twice in the steps that form the acyl enzyme intermediate and in the steps that break it down hydrolytically (Figure 16.3). The combination of histidine residue as general base and serine- OH as an incipient nucleophile has been so successful it has become a dominant paradigm for peptide bond cleaving catalytic architectures. 6

16.3.2 Histidine as a Nucleophile: Histidine Autokinases

Remarkably efficient and in widespread use as histidine imidazole side chains are as general base-acid catalysts noted above, the employment of the abundant neutral imidazole ring form as nucleophile is dominant in the thousands of members of the bacterial sensor kinases by which bacteria sample their local external environments. 7

The histidine kinases are the first component in two-component regulatory systems that sense an external physical or chemical condition or specific molecule. These transmembrane sensor kinases relay that information inside the cell to turn on or off specific sets of genes that encode a metabolic or behavioral response (Figure 16.4A). Some bacteria express 50 or more two-component systems at the same time to integrate external inputs. 8

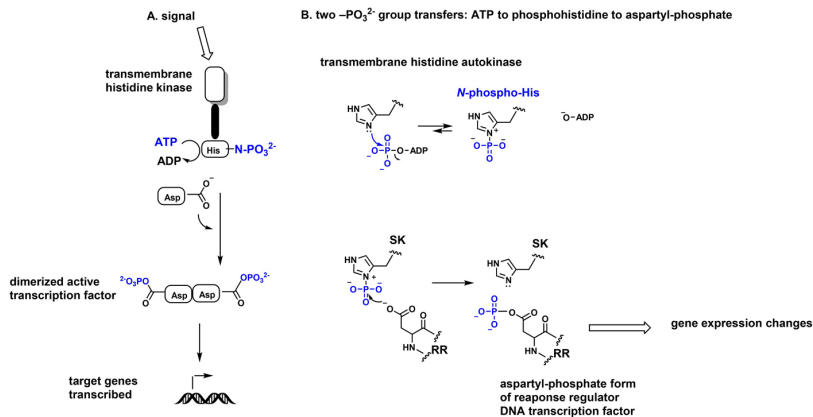


Figure 16.4 Bacterial protein histidine autokinases in two component systems. (A) Schematic for transmission of an external signal to target gene activation by transmembrane signaling. Sensor histidine kinase and DNA transcription factor response regulator. (B) Two phosphoryl transfer steps: ATP to an N -phosphohistidyl residue on a kinase to an aspartyl phosphate group on a response regulator, converting it to the active DNA binding dimer.

The second protein components of two-component systems are termed response regulators. Prototypically, these are two-domain proteins. The first domain has an aspartyl side chain that receives the phosphoryl group from the autophosphorylated histidine kinase. The second domain is usually a transcription factor domain that achieves its active DNA binding conformation when the receiver domain has its aspartate phosphorylated (Figure 16.4B), often by induced dimerization.

Figure 16.4B shows that the two-component system chemical logic involves two successive phosphoryl group transfers. The first is attack of the imidazole nitrogen on P_γ of ATP, constituting the sensor kinase autophosphorylation step in response to recognition of a signal by the external domain of this transmembrane sensor kinase. The second phosphoryl group transfer is from that phospho-histidine *phosphoramidate* to acyl phosphate in the aspartyl phosphate side chain of the response regulator protein. Aspartyl phosphates are both thermodynamically activated and kinetically labile to hydrolysis. Such hydrolysis is an autonomous clock that turns off the transcription factor, the expression of target genes, and ultimately the metabolic and/or behavioral response to the external signal.

Across bacterial species there are now thousands of histidine kinases known: two-component systems dominate bacterial kinase chemical biology. 9 During evolution of eukaryotic signaling logic, histidine kinases as transmembrane sensors have largely been replaced by transmembrane tyrosine autokinases and paired cytoplasmic tyrosine or serine kinase partners. However, in soluble metabolic enzymes, such as nucleoside diphosphokinase and succinyl coenzyme A (CoA) synthase, transient *N*-phosphohistidinyl enzyme intermediates form in each catalytic cycle (Figure 16.5). Likewise, the bacterial phosphotransferase system for transport of hexoses into cells concomitant with phosphorylation involves a cascade of transient P-His phosphoramidate protein intermediates. 10

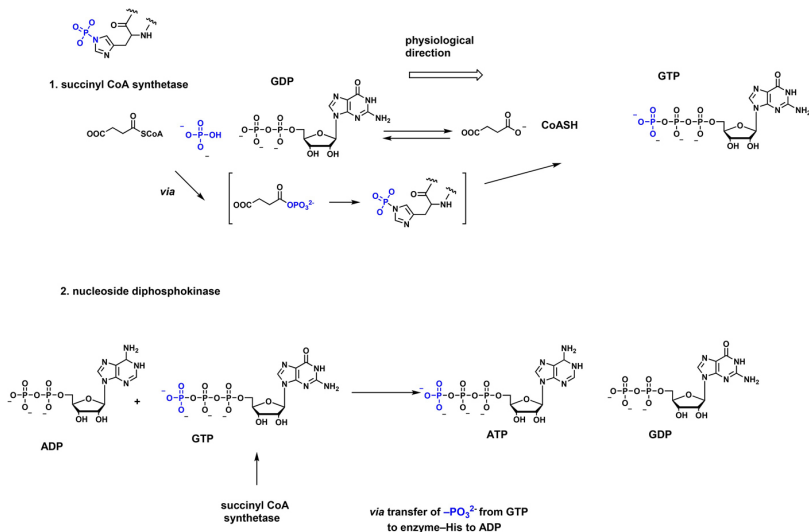


Figure 16.5 Two mammalian metabolic enzymes that catalyze net phosphoryl group transfers by way of *N*-phosphohistidinyl-enzyme adducts: succinyl CoA synthetase and nucleoside diphosphokinase.

In general, the $N-PO_3^{2-}$ phosphoramidate group in either N_1 or N_3 -phosphohistidine are not stable to hydrolytic decomposition to inorganic phosphate

and histidine. The N-PO_3^{2-} bonds retain the high thermodynamic phosphoryl group transfer potential found in ATP side chain phosphoric anhydrides. That activation accounts for the two-component systems' abilities to transfer the phosphoryl group from P-His to aspartate side chains in partner response regulator–transcription factors. In turn, very few stable P-His covalent adducts have been observed in posttranslational phosphorylation of proteins. P-His residues are generally about catalysis, not stable allosteric protein posttranslational phosphorylations.

16.3.3 Posttranslational Methylation at N_1 or N_3 of the Imidazole Ring

The above examples of histidine *N*-phosphorylations occur as transient covalent enzyme adducts during catalytic turnover events, any one of which lasts less than one second. In contrast, there are rare proteins where covalent *N*-methylation of the imidazole ring side chain of a histidine residue is a stable, long-lived modification. ¹¹ The classic case has been histidine₇₃ in actin. ¹² Surprisingly, two distinct N_1 -methyl and N_3 -methyl imidazole regioisomers of His₇₃ are found, produced by action of two distinct *S*-adenosylmethionine-dependent methyltransferases (Figure 16.6). More recently, histidine *N*-methylations have been detected in myosin, in proteins that make up mitochondrial complex I, and in yeast L3 protein of ribosomes, required for subunit assembly. ¹³ The paucity of histidine side chain posttranslational methylations contrasts with the abundance of lysine N_6 methylations, particularly in histone proteins H3 and H4 noted in the next section.

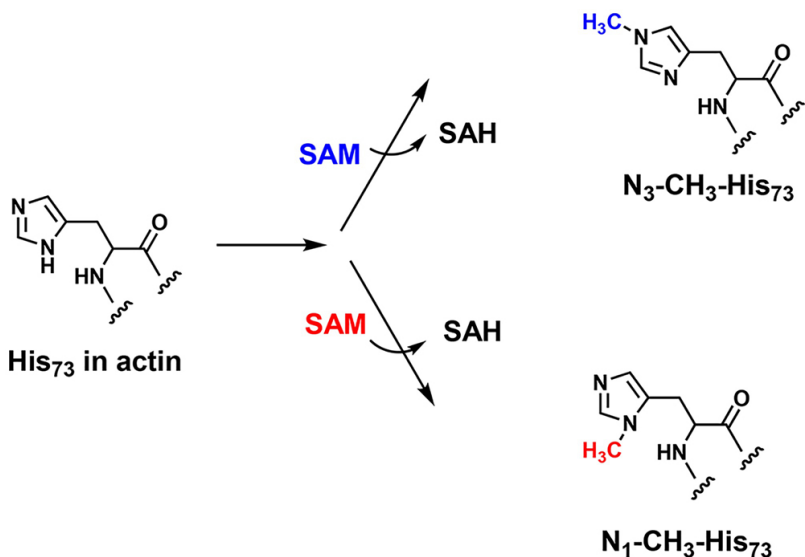


Figure 16.6 Two distinct histidine *N*-methyltransferases with distinct regiospecificity for N_1 or N_3 of the imidazole side chain of histidine₇₃ in actin.

16.4 Protein Lysine N_6 as a Nucleophile for Alkylations, Acylations, AMPylations and

Imine–Enamine Tautomerizations

Of the side chain nitrogen groups in proteins, lysine N₆ amino groups undergo the broadest range of covalent modifications by electrophiles other than protons. Iminium ion formation with pyridoxal-phosphate (PLP) converts apo forms of PLP–enzymes into the coenzyme-bound holo forms ready for a range of amino acid-based chemistry at C₂, C₃ or C₄ of the substrates as detailed in previous chapters. A distinct catalytic context for Lys N₆ iminium ion formation occurs with ketone substrates to set up low energy aldol cleavage chemistry.

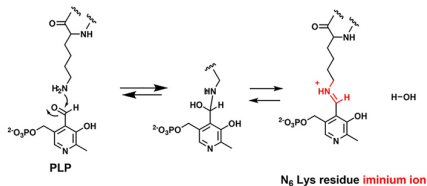
Lysine side chains undergo one-carbon alkylations *via* methyl group transfer from S-adenosyl methionine (SAM) once, twice or three times, especially at N-terminal tails of histone residues. A wide range of lysine side chain protein acylations occur, with several acyl groups from abundant cellular acyl CoAs. Finally, acylation of lysine side chain nitrogens is not limited to small (C₂–C₁₆) acyl chains. Several hundred protein acylations with the information-rich, small (75-residue) folded protein ubiquitin control trafficking and signaling across many eukaryotic metabolic pathways. Analogously, isoforms of the approximately 100 amino acid residue small ubiquitin-like modifier (SUMO) proteins are structural analogs of ubiquitin and undergo catalyzed acylation of lysine side chains on target proteins in mammalian cell biology.

16.4.1 Iminium Ion Formations

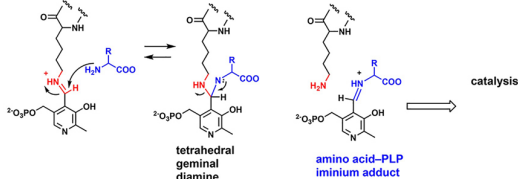
Active site lysine N₆ nitrogens in their deprotonated amino form are the nucleophiles that convert approximately 100 inactive apoenzyme forms to active holo forms of pyridoxal phosphate enzymes involved in all facets of amino acid metabolism. [14•15](#) The general strategy is to tether the PLP aldehyde iminium ion linkage in the resting form of all PLP enzymes ([Figure 16.7A](#)). These are effectively priming reactions to create catalytically competent holoenzymes. There are essentially no free untethered PLP molecules inside cells. As discussed in Chapter 6, Section 6.3 on transaminases, and in other chapters, these resting enzyme–Lys–N₆=PLP imines undergo transaldimination with the amino group of an amino acid to initiate every catalytic cycle ([Figure 16.7A](#)) and then to end every catalytic cycle.

A. lysine residue N_6NH_2 : imine formation with PLP

1. inactive apo enzyme to active holo enzyme

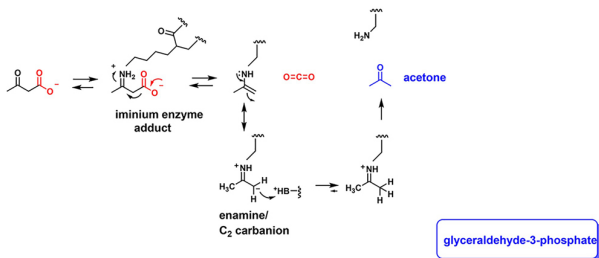


2. transaldimination to start each catalytic cycle



B.

1. acetoacetate decarboxylase



2. iminium formation with substrate carbonyl for C-C bond formation or cleavage

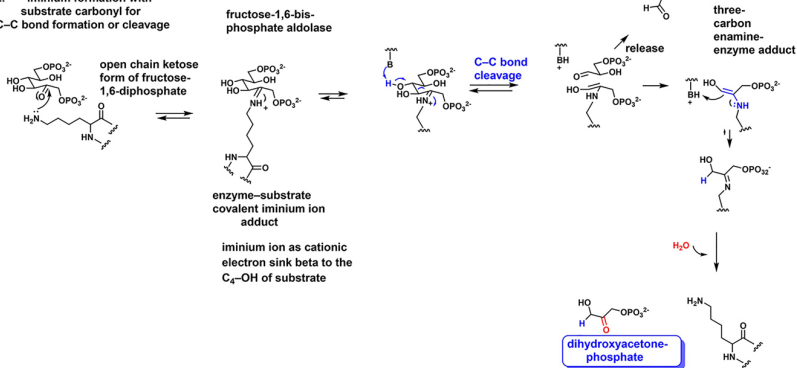


Figure 16.7 Imine formations between lysine residue side chain N_6-NH_2 . (A) (1) Tethering of coenzyme pyridoxal-phosphate to active site lysine to generate the active forms of approximately 100 PLP-enzymes. (2) Transaldimination with amine acid substrate creates aminoacyl-PLP aldimine as first step in catalyses. (B) (1) Iminium formation between acetoacetate and a lysine residue in acetoacetate decarboxylase sets up an iminium ion sink beta to the substrate carboxylate, lowering energy for decarboxylation. (2) Iminium ion-enamine catalysis in the action of fructose-bis phosphate aldolase in either C-C bond cleavage direction (shown) or in C-C bond-forming condensation reaction during gluconeogenesis (not shown).

A distinct context for side chain lysine- N_6 iminium ions occurs with linkages to carbonyl-containing substrates. Two examples are bacterial acetoacetate decarboxylase 16 and fructose-1,6-bisphosphate (FBP) aldolase 17 (Figure 16.7B). The former enzyme was used at an industrial scale in world war I for production of the solvent acetone, while the latter enzyme is a central catalyst for breakdown of hexoses in glycolysis, and diurnally in the reverse direction the process of gluconeogenesis.

In both enzymes the facile imine to enamine tautomerization brings an accessible stabilized carbanion equivalent into play to effect C–C bond cleavage. In the gluconeogenesis back direction, the FBP aldolase enables C–C bond formation between two three-carbon fragments. Lys- N_6 iminium ion formation with the C_3 ketone of acetoacetate sets up an iminium ion as an electron sink beta to the carboxylate ion to be eliminated as CO_2 .

In acetoacetate decarboxylase that iminium ion electron sink lowers the energy barrier for C–C bond cleavage because the resultant C_2 carbanion is delocalized as the enamine. Since nitrogen is more electronegative than carbon the predominant contributor to the C_2 carbanion structure is the neutral enamine. That can protonate at C_2 to form the product acetone iminium adduct. Hydrolysis releases acetone.

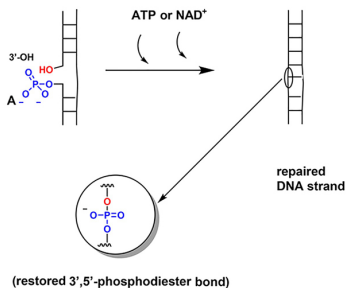
Similar logic applies to FBP aldolases that use comparable iminium ion chemistry as an active site Lys N_6 amine attacks the C_2 ketone of the open chain form of FBP. That cationic iminium ion is beta to the C_4 –OH of the fructose framework, set up for a low energy aldol cleavage, breaking the C_3 – C_4 bond of the hexose into two three-carbon fragments. Glyceraldehyde-3-phosphate is released from what had been carbons 4–6 of FBP. Carbons 1–3 are still tethered to the enzyme as the enamine or stabilized carbanion. Protonation of that carbanion at C_3 yields the iminium adduct of dihydroxyacetone-phosphate. Hydrolysis of the imine linkage regenerates the Lys- N_6 amine residue for the next catalytic cycle while releasing free dihydroxyacetone-phosphate.

These two C–C cleaving enzymes illustrate the catalytic power of iminium ion–enamine tautomerization to lower energy barriers for carbanion equivalents in enzymatic catalysis. Recall that, as described in Chapter 10, thiamin-PP, the coenzyme form of vitamin B_1 uses similar β -iminium electron sink logic to facilitate comparable substrate stabilized carbanionic intermediates.

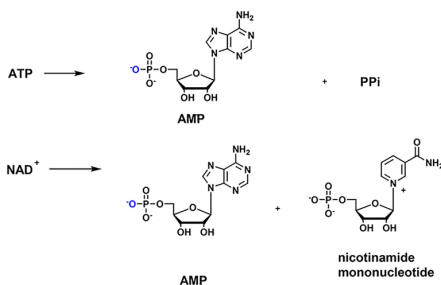
16.4.2 AMPylations in DNA Ligase Catalytic Repair of DNA Strand Breaks

DNA ligases act both in normal DNA replication by ligating discontinuous Okazaki fragments on the lagging DNA strand left during DNA polymerase action and in repair of both single strand nicks and double strand breaks of DNA double helical templates. An estimated 200 000 single strand breaks and approximately 50 double strand breaks in each human cell's three billion base pair genome occur every day and must be repaired. DNA ligase isoforms are busy as they convert gaps presenting 3–OH and 5'–phosphate ends for DNA strand repair of the covalent bridging 3',5'-phosphodiester bridge (Figure 16.8A).

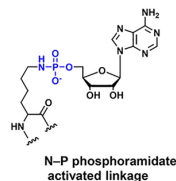
A. DNA: ligase; acting on single strand nicks



energy input and net cleavage stoichiometry



Covalent enzyme-Lys N_6 -AMP intermediate
as adenylyl transfer reagent to 5'-side of DNA gap



B. structure of the covalent
Lys-AMP phosphoramidate
adduct in DNA ligase

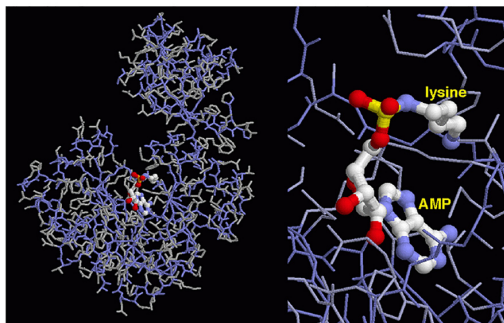
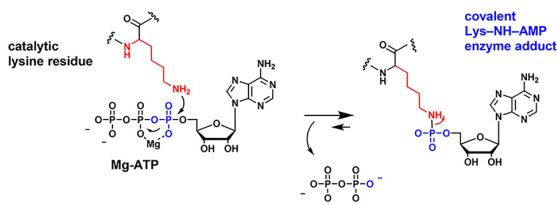


Figure 16.8 Enzyme-lysine-N-AMP adduct as intermediate in DNA ligase-mediated rejoining of gaps in cleaved DNA strands. (A) Either ATP or NAD^+ may serve as an AMP moiety donor in different organisms. (B) View of the covalent ligase-Lys-AMP adduct, from PDB 101 molecule of the month.

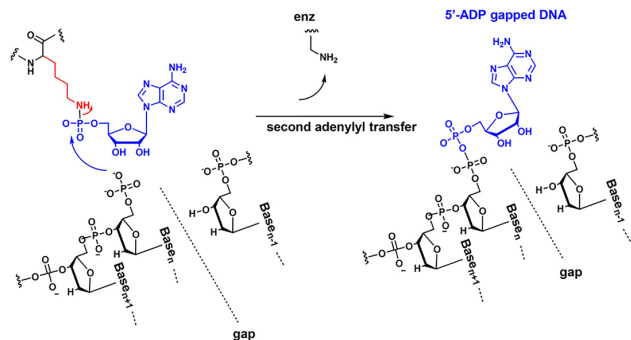
The ligases carry out a three-step repair mechanism, utilizing two adenylyl transfers 18 (AMP moiety) (Figure 16.9). DNA phosphodiester bond ligation requires energy input, usually in the form of ATP side chain cleavage but in some organisms NAD^+ cleavage is the replacement strategy. In either case, an active site lysine residue acts as an initiating nucleophile to form a covalent lysyl- N_6 -AMP phosphoramidate (N-P bond) linkage (Figure 16.9), releasing PPi from ATP or nicotinamide mononucleotide (NMN) from NAD^+ . A view of the enzyme-Lys-AMP adduct in the ligase active site is depicted in Figure 16.8B.

DNA ligase action three step mechanism

1. Formation of enzyme-Lys-N₆-AMP covalent intermediate: first adenylyl transfer



2. transfer of AMP moiety from active site lysine to 5'-P side of the DNA gap, generating a 5'-ADP-gapped DNA strand



3. repair: formation of 3',5'-phosphodiester connecting bond to bridge the single strand break

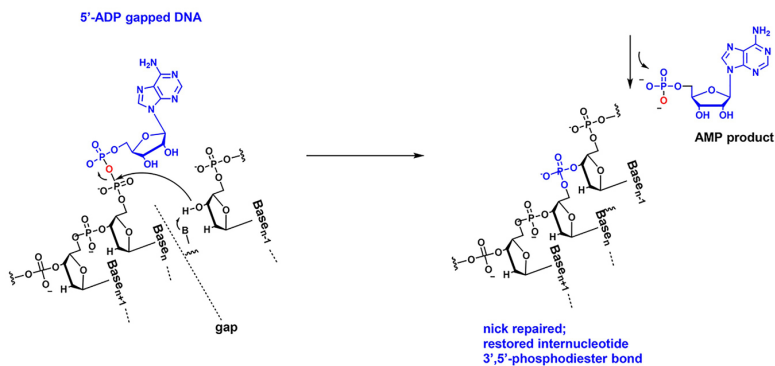


Figure 16.9 Consecutive transfers of the AMP moiety: (1) ATP to enzyme-Lys-N-AMP; (2) AMP transfer to oxanion of 5-phosphate on one side of the gapped DNA; (3) attack by the 3 OH on the other side of the gap restores the internucleotide phosphodiester bond and releases free AMP.

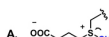
The Lys-AMP covalent enzyme adduct then acts as transfer agent for the electrophilic AMP group in the second step of the ligase reaction. The attacking nucleophile is one of the oxanions on the 5'-phosphate side of the DNA strand gap (Figure 16.9). This regenerates the free active site lysine side chain while forming a 5'-ADP ester on the broken DNA strand. The third step of the ligase reaction is the actual

repair step. The 3'-OH on one side of the gap attacks the internal phosphoryl group of the ADP moiety on the 5'-side of the gap. This reforms the bridging 3' to 5' phosphodiester in repaired DNA as AMP departs.

Overall, the energy released in hydrolysis of the phosphoric anhydride bond of ATP (to AMP and PPi) has been utilized to drive DNA gap repair *via* phosphodiester bond reformation. Mechanistically, the 5'-phosphate monoester side of the gap has to be activated for capture by the 3'-OH, meaning one of its PO₃ oxygens is derivatized to lower its energy as a leaving group. That O⁻ atom has become part of an AMP-O moiety as it departs (pK_a of HOH to OH⁻ = 14, pK_a of AMP-OH to AMP-O⁻ = 6, so elimination of AMP is up to 10⁸ times more favorable). To get to that activated 5'-ADP side of the DNA gap, the enzyme uses a covalent Lys-N-AMP intermediate as the proximal adenylyl transfer reagent to the gapped DNA. *Inter alia*, it reveals that the Lys-N-AMP is a high group transfer adenylyl transfer reagent, used millions to billions of times per cell per day. The initial adenylyl transfers from ATP to the active site lysine residues have been termed AMPylations in the recent literature and occur in other enzyme cycles. These are real but fleeting catalytic intermediates, not stable posttranslational modifications of client proteins.

16.4.3 Iterative Methylations

There are some 70 human methyltransferases that act to add up to three [CH₃ +] equivalents to protein side chains (Figure 16.10). Lysine and arginine side chain nitrogens predominate, although a few histone imidazole side chain methylations are noted in the previous sections. ¹⁹ These are stable posttranslational modifications, rather than catalytic intermediates in enzymatic reaction cycles. The array of methylated proteins constitutes the methylome. These *N*-alkyl adducts are not reversed by direct hydrolysis. In fact, oxidative demethylation strategies are required to undo these posttranslational modifications. Families of methyl lysine oxidative demethylases (KDMs in Figure 16.10B) act in opposition to the protein lysine methyltransferases.



B.

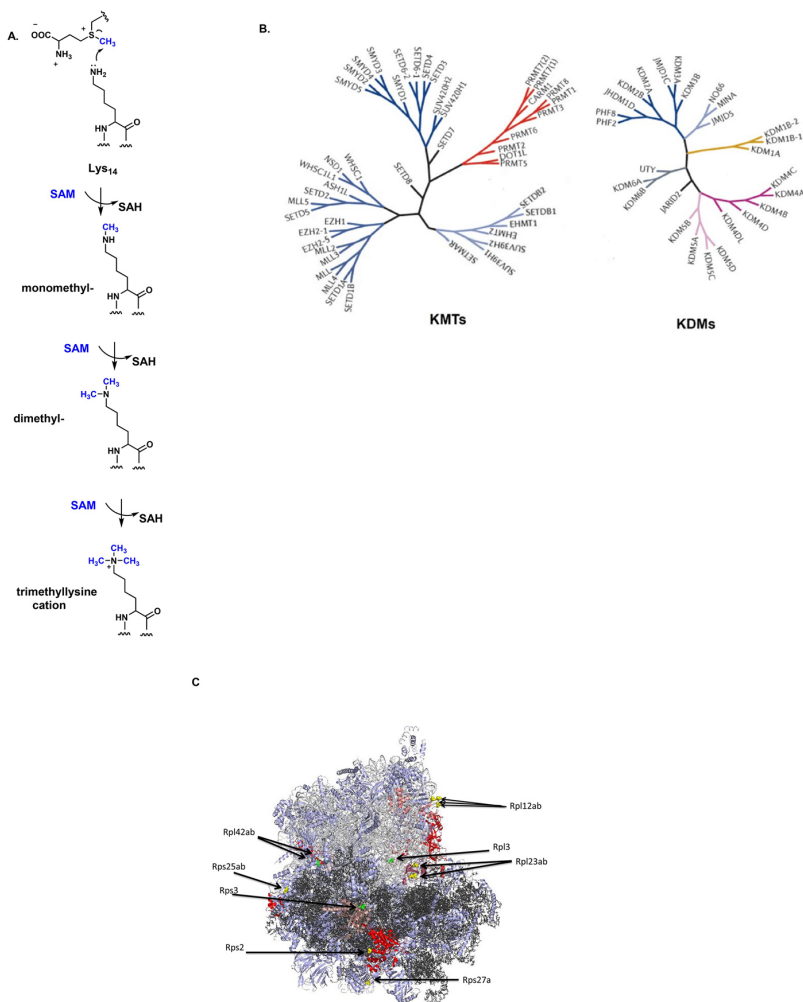


Figure 16.10 (A) Successive *N*-methylations from *S*-adenosylmethionine to the *N*₆-amine side chain of Lys₁₄ in ribulose-bis-phosphate carboxylase create the active *N,N,N*-trimethyl-lysine cationic form of this key plant enzyme in carbon fixation. (B) Family tree diagrams for human protein lysine methyltransferase (KMTs) and for the demethylases (KMDs) that oppose the methyltransferase by oxidative deaminations to reverse the *N*-methyl-Lys residue covalent marks. Reproduced from [ref. 58](#) with permission from Bentham Science Publishers Ltd, Copyright Eureka Science. (C) View of yeast ribosome with proteins from large (Rpl) and small subunits (Rps) that become methylated on *N*₆ of Lys side chains. Reproduced from [ref. 21](#) with permission from Elsevier, Copyright 2013.

While the *N*-terminal tails of histones H3 and H4 have garnered the most attention over the past twenty years, other proteins are also part of the methylomes of different organisms. In fact, the first identified protein methyltransferase was shown to add three methyl groups consecutively, from three molecules of SAM cosubstrates, to *N*₆ of Lys₁₄ of the large subunit of plant ribulose-1,5-bis-phosphate carboxylase, [20](#) the key enzyme in plant fixation of CO₂ globally. The trimethyl lysine 14 is now both cationic and hydrophobic ([Figure 16.10A](#)). Shown in [Figure 16.10C](#) is a view of the yeast

ribosome where four small ribosomal subunit proteins (Rps2, Rps3, Rps25ab and Rps27a) and four large subunit proteins are each methylated on N₆ of a lysine side chain. 21

The several members of the protein lysyl transferases that act on histone tails can add one, two or three methyl groups to particular lysine side chains (Figure 16.11) with different consequences for transcriptional activation or repression of genes in DNA that is in contact with those histones in nucleosomes. Histone posttranslational modification biology is about combinatorics of both methylations of Lys and Arg residues and acetylations, allowing for thousands of distinct molecular variants with potential and actual functional distinctions. The topic of posttranslational modifications of histones and effects on chromatin molecular biology could fill several volumes on gene regulation and is beyond the narrow focus here on lysine N-methylation posttranslational modifications.

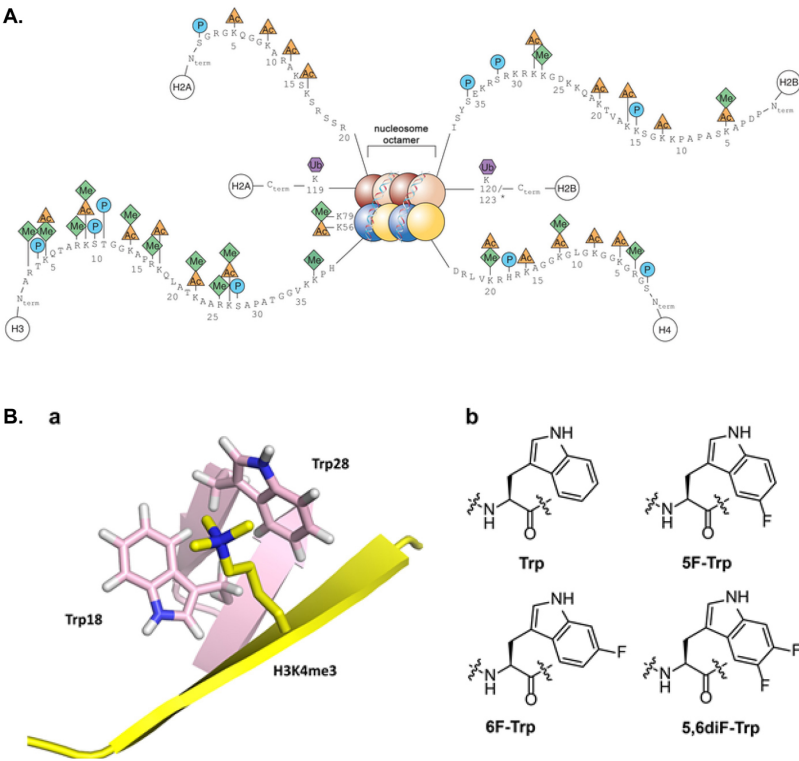


Figure 16.11 (A) Schematic of the possible posttranslational acetylations, methylations and phosphorylations of the N-terminal (unfolded) tails of the four histones that occur in pairs to constitute the nucleosome core that wraps around 143 nucleotides of DNA to compact the DNA genomes to fit within cells. Reproduced from ref. 22 with permission from Taylor & Francis, Copyright 2009. Acetylations and methylations can compete with each other for certain lysine side chains on H3 and H4 with opposite effects on gene transcription (deactivation vs. activation). H3 is the most heavily modifiable, H2A the least modifiable. (B) Diagram of a pair of tryptophans on a partner protein interacting with the cationic trimethyl lysine tail of a modified histone H3 region to enable recruitment of the partner protein. Reproduced from ref. 23, <https://doi.org/10.1038/s42004-020-0313-2>, under the terms of the CC BY 4.0 license, <http://>

Inspection (Figure 16.11) of the four histones (H2A, H2B, H3 and H4) in two copies each that comprise the protein components of eukaryotic nucleosomes, wrapping around 143 base pairs of DNA, shows a plethora of differential methylations, acetylations and phosphorylations of the four histones. 22 Methylation of Lys N₆ and the Arg guanidinium group are biased for the extended NH₂ tails of H3 and H4. Acetylations, to be discussed in a separate section, are more evenly distributed. The figure shows that acetylation or methylation can occur and compete with each other at certain lysine side chains, leading to different functions (gene transcriptional activation vs. repression). In the first 37 residues of histone H3 seven lysine side chains can be methylated. Only two lysines are methylated in the H4 tail, two in the H2B tail and none in the H2A tail, reflecting discrimination by the family of KMTs (Figure 16.10B) that act on accessible histone N-terminal tails.

As in the diagram of Figure 16.10A for trimethylation of lysine-14 N₆ nitrogen, histone tail methylations can stop at monomethyl or dimethyl or proceed all the way to the trimethyl permanently charged hydrophobic Me₃-Lys state. That hydrophobic positively charged cationic Me₃⁺N₆ group is the signal for attracting partner transcription complex proteins that have electron rich aromatic cages. Figure 16.11B depicts such a ditryptophan cage from a protein finger region of the trimethyllysine demethylase KDM5A interacting with the H3K4-Me₃ residue that is a robust signal for transcriptional activation. 23 The figure shows the virtue of adding three methyl groups to a lysine side chain to create a unique chemical recognition element.

16.4.3.1 *N*-Methylation of Peptide Bond Nitrogens

As an aside before turning back to lysine side chain acylations in proteins, some nonribosomal peptide synthetases (NRPS) assembly lines contain hybrid adenylation-*N*-methyltransferase catalytic domains. 24-25 After installation of the selected amino acid in thioester linkage to peptidyl carrier protein (PCP *n*) by the adenylation domain, the embedded *N*-methyltransferase catalyzes nucleophilic attack of the aminoacyl-S-PCP *n* on the methyl group of SAM. The result is the *N*-methylaminoacyl-S-PCP *n*. Subsequent coupling by the module's condensation domain with the upstream peptidyl-S-PCP *n*-1 generates an *N*-methylated peptide bond in the growing peptidyl-S-PCP *n* chain.

A more impressive peptide *N*-methylation catalyst is the *N*-methyltransferase from the ribosomally generated fungal peptide omphalotin A from the fungus *Omphalotus olearis*. Omphalotin A is a cyclic decapeptide, derived from proteolysis of a ribosomally generated protein precursor, where nine of the peptide bonds are *N*-methylated. The S-adenosylmethionine *N*-methyltransferase catalyzes *N*-methylation of the peptide NH bonds. 26 This is a surprising accomplishment, not least because the peptide nitrogens are poor nucleophiles so the chemical task is arduous. How the enzyme activates each of the nine peptide-amide bond nitrogens to achieve heightened nucleophilicity (twisting the peptide bonds out of planarity?) is unclear. A further anomaly is that the omphalotin core peptide precursor is covalently tethered to the *N*-methyltransferase domain. Proteolysis to carve out the methylated decapeptide happens after the peptide methylation steps. There is still much to learn about how proteins control amine or amide nucleophilicity in *N*-methylations.

16.4.4 (Reversible) Posttranslational Acylations by Acyl CoA Metabolites

Acylations of amines to form amides are the bedrock of amine chemical biology, as

iterated in the several chapters of Section II on organic small molecule nitrogen metabolism. It is no surprise that amide formation by acylation of lysine N₆ amine side chains in proteins continues that theme in macromolecular as well as low molecular weight amines. The common acyl donors in low molecular weight metabolites are the acyl-S-CoA species, thermodynamically activated but kinetically stable in the enzymes of specific acyl transfer enzyme catalysts. Most notable is the two-carbon acetyl donor acetyl CoA, followed by the three-carbon branched malonyl acyl group and the four-carbon succinyl CoA (Figure 16.12A). The most abundant products of fatty acid synthases are the straight chain fatty acids C14 (myristoyl CoA), C16 (palmitoyl CoA) and in some organisms C18 (stearoyl CoA) acyl thioesters. Any of these activated acyl groups can be put to use for lysine side chain aminoacylations.

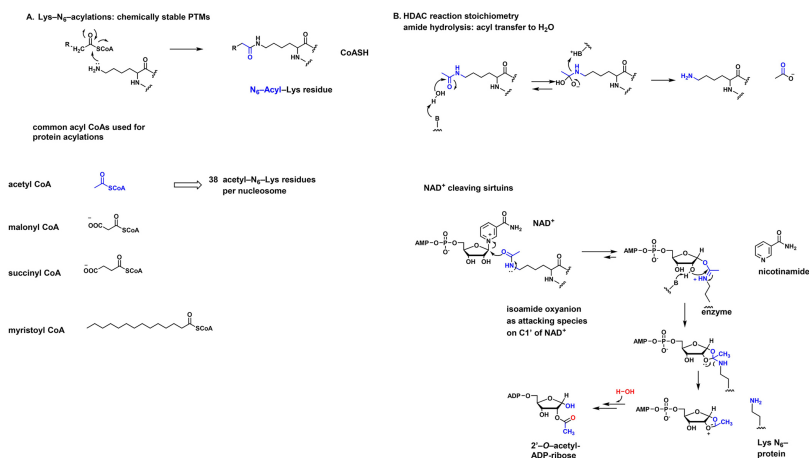


Figure 16.12 (A) Acyl groups in acyl CoAs that serve as substrates for posttranslational acylation of side chain amines or N-termini of proteins. (B) The actions of histone acetyl transferases (HATs) are reversed by histone acetyl decarboxylases. A subset of histone deacetylases require NAD⁺ as a cosubstrate and transfer the acetyl group to the ADP ribose moiety of NAD⁺. This, in turn, requires the acetyl fragment in flight to act as an initial nucleophile. 2'-O-acetyl-ADP ribose is the released product, not a straightforward hydrolysis of an acetyl group on a histone lysine amine side chain.

Figure 16.11A depicts sites of lysine N₆ side chain acylation on the four histone protein components of nucleosomes in addition to the methylation sites noted in the preceding section. 22 H2A has four lysines that can be acetylated (none are competitively N-methylated); H2B has four; five of the H3 lysines in the N-terminal tail can be either N-acetylated or N-methylated; H4 has five such acetylation sites. These total to $19 \times 2 = 38$ observed lysine N₆ acetylation sites in each nucleosome that carries two copies of each histone. This is high density posttranslational modification enzymology.

Several of the acetylations block competing methylations at those lysine residues, converting transcriptional activation signals to deactivation. Up to a dozen histone acetyltransferases are expressed in human tissues with distinct regiospecificities for particular histone tails, for particular lysine residues in the context of preexisting PTMs, and in the context of monomethylations vs. dimethylations vs. trimethylations.

The interplay of methylation, up to trimethyl-lysines, vs. acetylation has been described as writing a posttranslational code that then gets read by partner proteins,

leading to integrated signaling that results from repressive or activated protein complexes recruited to particular nucleosomes. In addition to the reading and writing the histone code(s) there is the erasing phase: enzymatic removal of the histone alkyl and acyl marks. 27-28

Reversals of other posttranslational modifications, such as phosphorylations, glycosylations and acyl-S-cysteinyl protein thioesters are usually hydrolytic. That is the case for the *N*-acetyl-N₆-lysine histone amide marks. As in the case of multiple histone acetyltransferases (HATs), there are multiple histone deacetylases (HDACs) that operate at some of the 38 possible *N*-acetyl groups on nucleosomes in different biological contexts. Also, while most HDACs are straightforward amide hydrolases (Figure 16.12B) there are some NAD⁺-dependent hydrolases that instead of direct hydrolysis transfer the acetyl group to the 2'-oxygen of the ADP ribose moiety of NAD to produce 2'-*O*-acetyl-ADP ribose. Clearly *N*-deacylation biology for histones and other *N*-acetylated proteins is complex and regulated. 29

In contrast, there is no straightforward chemical route for hydrolytic cleavage of *N*-alkyl groups. They are indefinitely stable in and inert to water. Instead, nature, having chosen to use protein *N*-methylation for regulatory signaling, needed to invent an oxidative deamination strategy to terminate those biological signals. Oxidative deamination means O₂ as cosubstrate, which in turn means either flavin adenine dinucleotide (FAD) 30 or more broadly redox active iron enzymes 31 at the active site of those oxidative deaminases. A family tree of nonheme mononuclear iron-based lysine residue oxidative deaminases is depicted in Figure 16.10B. Chemically, the CH₃-NH- side chain is converted to the HO-CH₂-NH- products. These are unstable tetrahedral carbinolamines, with two hetero atoms – oxygen and nitrogen – on the same carbon. They break down nonenzymatically and rapidly in aqueous solutions to formaldehyde and the parent amine (Figure 16.13).

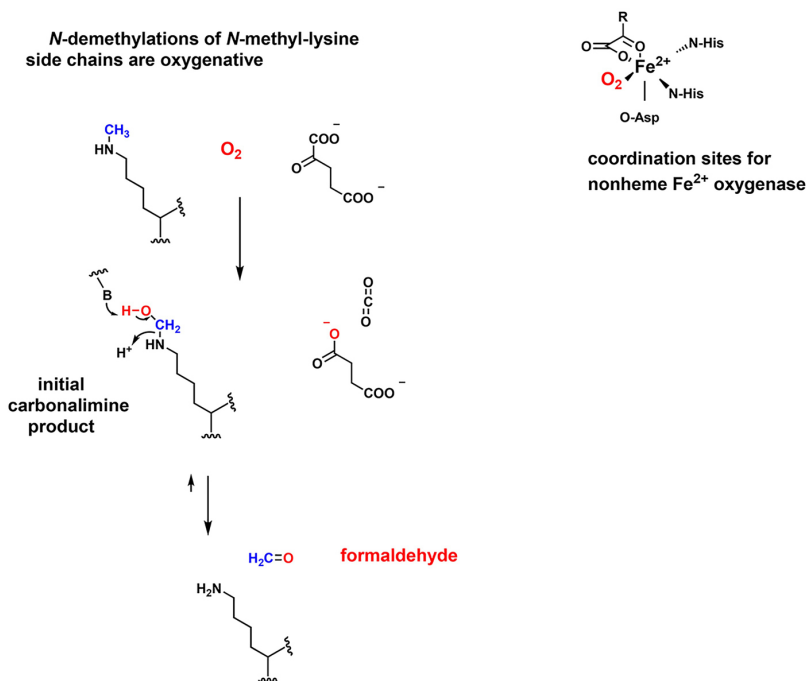


Figure 16.13 Oxygenative demethylation of *N*-methyl-lysine side chains.

A brief comment about long chain acyl amide formations in proteins. The paradigm is the *N*-myristoylation of some proteins with N-terminal glycine residues. ³² These are encoded as Met₁-Gly₂-proteins where the N-terminal methionine residue is removed cotranslationally by an N-terminal aminopeptidase. The newly liberated NH₂ of the now N-terminal Gly residue is then *N*-acylated by a myristoyltransferase, utilizing myristoyl CoA as an electrophilic acyl donor. Among the other properties imparted, the 14 carbon myristoyl group can serve as a subcellular address signal to associate the myristoylated protein with one or more subcellular membrane compartments. That isopeptide bond is not readily removed by the inventory of cellular aminopeptidases prior to complete proteolytic breakdown (Figure 16.14).

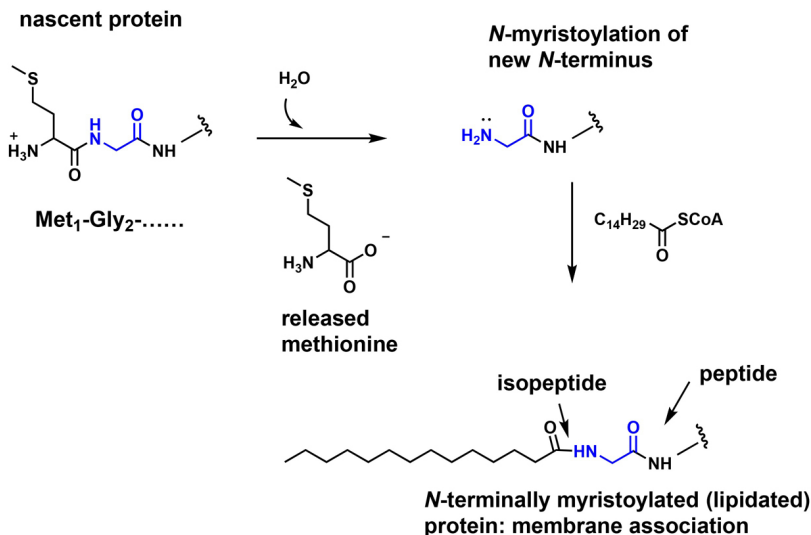


Figure 16.14 *N*-myristoylation of proteins can occur cotranslationally when a glycine is in the second position. Hydrolytic removal of the fMet₁ by an aminopeptidase uncovers a glycine as the new N-terminus. Its enzymatic *N*-myristoylation creates a membrane anchor for the lipidated protein.

16.4.5 Acylations by Information-rich Ubiquitin and SUMO Isoforms

Acylation on N₆-lysine side chains also occurs on hundreds to thousands of client proteins by acyl thioester acyl group donors where the acyl group is not a simple straight C₂-C₁₆ acyl chain but instead related classes of small 76–110 residue proteins that are folded and carry information. The classic information-rich protein in this family is ubiquitin ³³⁻³⁴ (Figure 16.15A), but a structurally related set of small ubiquitin-like modifier proteins ³⁵⁻³⁶ (SUMOs) (Figure 16.15A) are also widespread in eukaryotic biology. Both ubiquitin and SUMO have generated a monumentally large research literature as investigators have analyzed the hundreds of situations where ubiquitin or SUMO proteins have been added posttranslationally to client proteins to create isopeptide bonds. Those isopeptide bonds can then be removed hydrolytically once their signaling function has been achieved by sets of dedicated isopeptidases. We cannot attempt to cover the rich biology of these reversible protein tags in this short

section. Again, we focus only on the underlying chemical logic for lysine side chain acylation and diacylation as part of the continuum of lysine side chain reactivity.

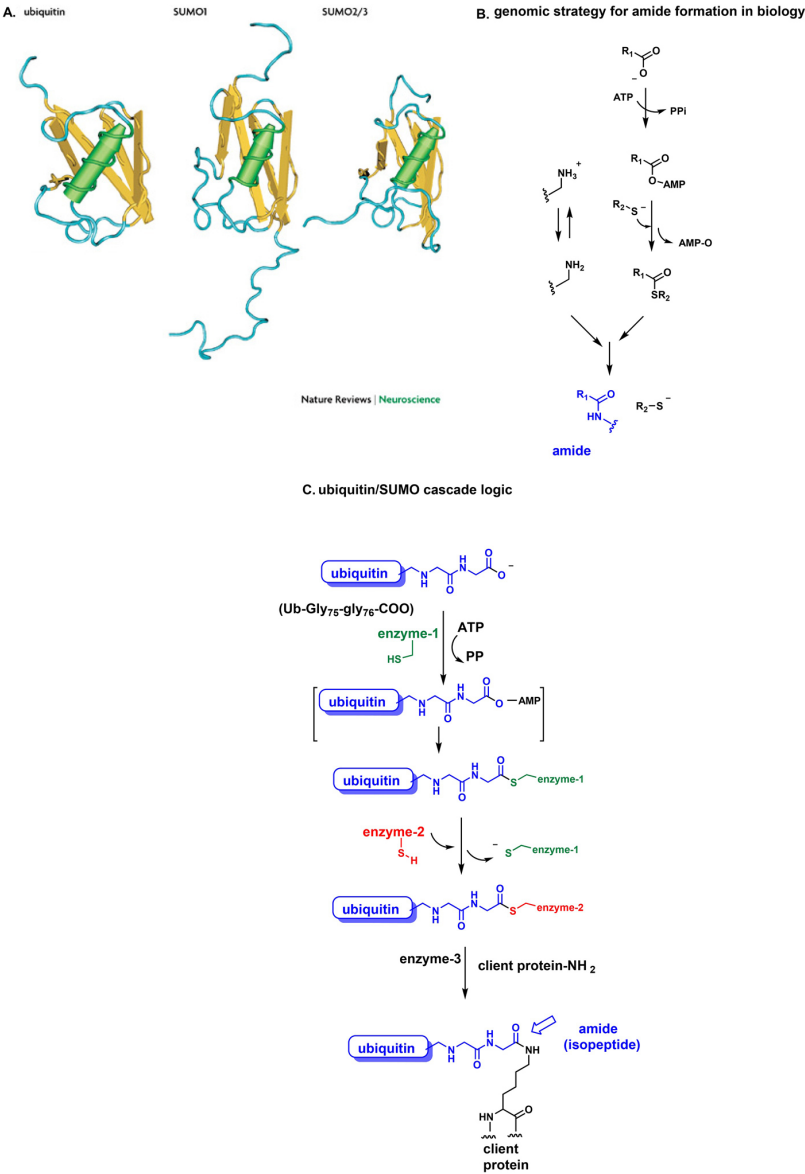


Figure 16.15 Ubiquitin and small ubiquitin like molecules (SUMOs). (A) Both ubiquitin and SUMOs are structurally related and serve as information-rich acyl tags for posttranslational acylations of lysine side chains in hundreds of client proteins. Reproduced from [ref. 59](#) with permission from Springer Nature, Copyright 1969. (B) A reminder of the general logic for amide bond formation (in this context *isopeptide bond* formations): ATP hydrolysis is required to activate the carboxylate, typically as the acyl-AMP. To gain kinetic stability while maintaining thermodynamic activation, the acyl group is transferred to a thiol group to

circulate within cells as acyl thioesters. (C). In the ubiquitin or SUMO context the acyl groups are information rich folded domains, the carboxylate is the terminal carboxylate of a Gly–Gly–COO[−] terminus, and the thiol is a cysteine side chain, first in enzyme-1 and then in enzyme-2. These thermodynamically activated ubiquitinyl- or SUMO–S–enzymes-2 are the acyl donors to lysine residue side chain amines on client proteins to form stable isopeptide bonds.

To form an amide bond between ubiquitin or SUMO isoforms and a lysine side chain in a client protein where ubiquitin or SUMOs are the acyl donors ([Figure 16.15B](#)) requires that ubiquitin or SUMO carboxylate groups be thermodynamically activated yet kinetically stable enough to diffuse to interact with specific client proteins. In previous cases of biological amide formations cells have used ATP to activate a carboxylate group as a mixed carboxylic phosphoric acyl-phosphate or acyl-AMP and then captured the acyl group with a thiolate to harness the thermodynamic acyl group transfer potential but in a kinetically stable acyl thioester.

That is exactly the chemical strategy translated from small molecule acyl CoAs as acyl donors to ubiquitin- and SUMO-thioesters. In the ubiquitin or SUMO cases the thiol is not coenzyme A but a sequential pair of active site thiolate proteins, known as enzymes 1 and 2 as noted in [Figure 16.15C](#). The carboxylate to be activated in both ubiquitin and SUMOs is at the C-terminus where a sterically unhindered –Gly–Gly–COO[−] terminus offers access to enzymes 1 and 2. For ubiquitin those glycines are residues 75 and 76 in the 76-residue folded protein. For SUMOs, that Gly–Gly motif is internal, four residues from the native C-terminus, and requires a dedicated protease to trim nascent SUMOs to release that tetrapeptide and create the Gly–Gly–COO[−] terminus for modification.

The first step, catalyzed by enzyme 1 in the ubiquitin cascade and a homologous enzyme in the SUMO cascade, catalyzes attack of the C-terminal carboxylate on P_{α} of cosubstrate ATP to release inorganic pyrophosphate and produce the ubiquitin–AMP or SUMO–AMP mixed anhydrides at the C-terminal glycine. These labile protein–adenylates are then captured by a thiolate anion of a cysteinyl residue in the active sites of enzymes-1, linking ubiquitin or SUMO in C-terminal thioester to their respective enzymes-1.

Typically, enzyme-1 in both ubiquitin and SUMO transfer cascades shows no interaction with client proteins. That is the job of enzyme-2 and a host of enzyme-3 architectures. Enzyme-2 transfers the activated ubiquitin or SUMO molecules from enzyme-1 to an enzyme-2 cysteine thiolate, creating covalent ubiquitinyl- or SUMO–S–enzyme-2 covalent adducts. The enzyme-2 proteins recognize specific partner proteins. In many cases the enzyme 3s are in effect adaptor proteins, physically connecting the ubiquitinyl-thioester or SUMO-thioester on enzyme-2 with a diverse range of client proteins, accounting for thousands of specific ubiquitin or SUMO transfers.

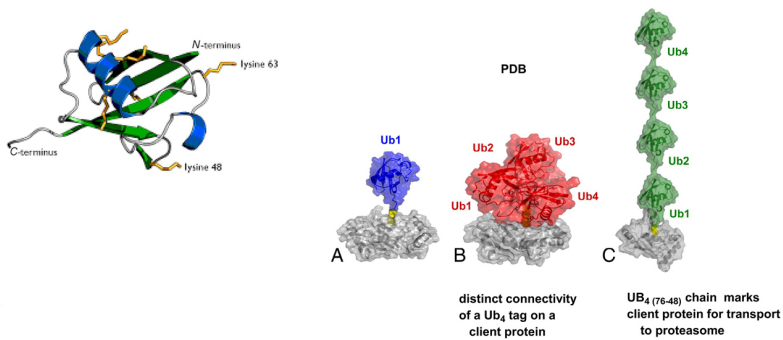
Almost all (>99%) of the transfers of the ubiquitin or SUMO thioesters to client proteins reflect capture by the N₆ amine of lysine side chains. Thus, almost all the covalent connections of ubiquitin or SUMO isoforms to client proteins are isopeptide or amide linkages.

The downstream biological consequences of monoubiquitin or SUMO isopeptide bonds affect almost every facet of eukaryotic cell biology and even a listing of functional changes is beyond the scope of this discourse. Simply put, the information-rich acylation of client proteins by ubiquitin and SUMOs is the dominant strategy for covalent marking of proteins in eukaryotic biology.

One clear functional distinction between ubiquitination and SUMOylation of client proteins is the oligomerization of ubiquitin chains attached to a client protein. Ubiquitin has seven lysine residues in its 76-residue polypeptide backbone. Each of them can act to capture an activated ubiquitinyl–S–enzyme-2 with or without the partnership of an enzyme-3.

Polyubiquitylation is the major strategy for marking tagged proteins for early hydrolytic destruction by the intracellular chambered protease nanomachinery of proteasomes. A minimum length of four covalently self-tethered ubiquitins (Figure 16.16A), typically *via* Lys48 to Gly76 inter-ubiquitin isopeptide bonds, as a tag on a client protein are sufficient to mark that protein (Figure 16.16A and B) for transport to the proteasome. Once there, ATPases spend multiple ATPs to unfold the client protein and feed its polypeptide backbone into the mouth of the chambered proteasome. 37•38 Inside the proteasome, digestion by inwardly facing proteases decomposes that client protein into small peptide fragments. Proteasome-mediated proteolysis of intracellular proteins is the main strategy for control of lifetimes of the thousands of distinct intracellular proteins, including digestion of proteins to peptides for presentation on immune receptors. SUMOs do not participate in comparable self-oligomerization or serve as covalent markers for accelerated degradation of their tagged client proteins. 35

A.



B. isopeptide bond hydrolysis by thiol-containing isopeptidase

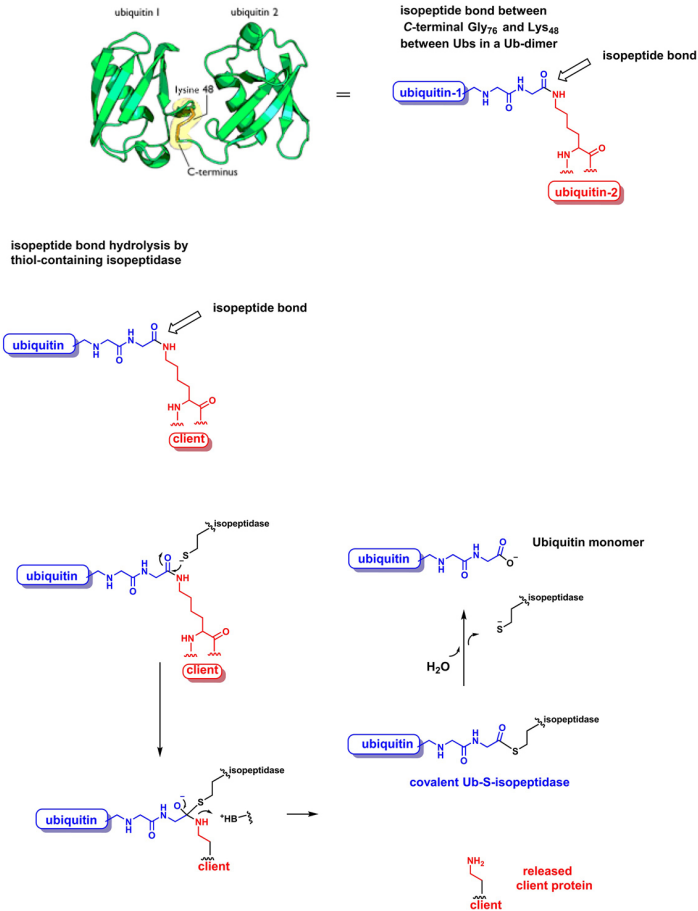


Figure 16.16 (A) Ubiquitins (Ubs), but not SUMOs, are often processed to poly-ubiquitin chains attached to client proteins. Shown is the structure of a Ub monomer and a Ub dimer linked *via* Lys₄₈ to the C-terminal Gly₇₆ of the neighboring Ub (also displayed is a space-filling model of a monoubiquitin adduct to a client protein and two different regioisomers of tetraubiquitin chains attached to a client protein). The extended Ub₄ regioisomer with K₄₈–G₇₆ isopeptide linkages is the most common Ub₄ regioisomer and signals for recruitment of the

Ub₄-tagged protein to be recruited to the proteasome for proteolytic destruction. Top left ubiquitin: Reproduced from <https://en.wikipedia.org/wiki/Ubiquitin>, under the terms of the CC BY-SA 3.0 license, credit: Rogerdodd. Top right: Reproduced from ref. 60 with permission from PNAS. Bottom ubiquitin: Reproduced from <https://en.wikipedia.org/wiki/Ubiquitin>, under the terms of the CC BY-SA 3.0 license, credit: Rogerdodd. (B) Isopeptide-specific deubiquitinases can cleave the linkages between ubiquitin monomers or between ubiquitin and client proteins. These hydrolases reverse the biological effects of protein tagging and/or retrieve the folded Ub tags before a client protein is fed into the proteasomes for degradative hydrolyses.

Because ubiquitin and SUMO covalent tags to client proteins are isopeptide bonds they are resistant to the standard inventory of cellular proteases that cleave normal peptide bonds. Instead there are dedicated isopeptidases encoded by and expressed in cells capable of hydrolyzing the isopeptide bonds between ubiquitin 39 or SUMO 40 and lysine side chains of client proteins (Figure 16.16B). The isopeptidases act to reverse the covalent tags, enabling both ubiquitin and SUMO tagging to be reversible signaling events, adding to their utility in the cellular signaling armamentariums. They also allow recycling of ubiquitin and SUMO without having to spend the energy for *de novo* biosynthesis. For example, the oligoubiquitin or polyubiquitin tags are typically cleaved from client proteins by isopeptidase action before the client protein is unfolded and fed into the proteasome, thus conserving folded ubiquitin oligomers that can be hydrolyzed down to the monomeric ubiquitins for reuse by enzyme-1.

Organisms have taken advantage of the nucleophilicity of the N₆-NH₂ of lysine residue side chains to evolve a plethora of enzyme-catalyzed posttranslational modifications, from short chain acetylations of histones to the enormous web of information in the proteins ubiquitin and SUMOs to control temporal and spatial facets of cell biology.

16.5 Arginine, an Unlikely Protein Side Chain Nucleophile

With a pK_a value of 12.5 for the guanidinium group of free arginine, only one part in 10^5 is the neutral guanidine group at pH 7.5. Since the positive charge is delocalized over both terminal nitrogens of the guanidinium group, both have their lone pair of electrons partially tied up with distinctly lowered nucleophilicity. On that basis arginine side chains in proteins, assuming essentially equivalent guanidinium group pK_a s, are unlikely candidates for either acid-base catalysis or nucleophilic catalysis. On the other hand, the bidentate positive charge is aptly dispersed for them to serve as chelators for dianionic phosphate groups that have become covalently attached during protein posttranslational phosphorylations. The resultant electrostatic interactions stabilize phosphoprotein conformations (Figure 16.17).

bidentate chelation with: 1. carboxylate side chains

2. phospho-Ser or phospho-Thr side chains

3. Tyr-O-sulfate side chains

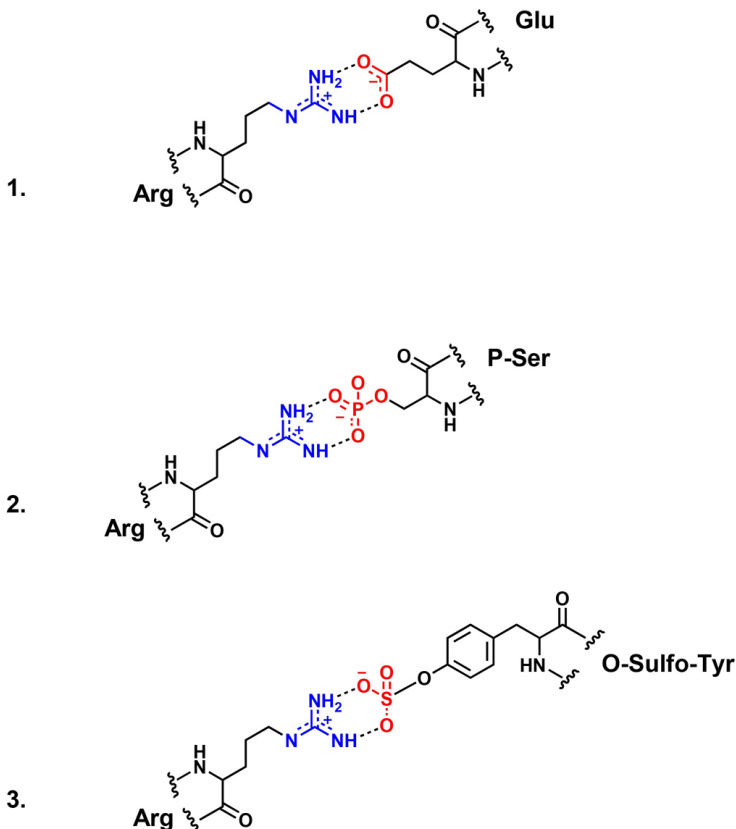
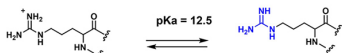


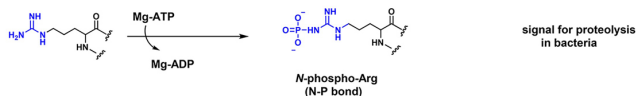
Figure 16.17 Arginine residue guanidinium cationic side chains serve as bidentate electrostatic ligands in ion pairs with carboxylate side chains (Asp or Glu) phosphomonoester side chains (P-Ser or P-Thr) and the rare sulfotyrosine monoanionic side chains within and between proteins.

Nonetheless, there are four types of posttranslational modifications, and possibly five, where the guanidinium group of arginine side chains in proteins has been covalently modified by an electrophilic fragment of a small molecule cosubstrate. [41](#) There is no redox evident in those group transfers so the guanidinium group by default has acted as a nucleophile, perhaps by deprotonation in the microenvironment of the posttranslational enzyme catalyst. [Figure 16.18](#) summarizes the five observed posttranslational modifications on arginine guanidino groups.

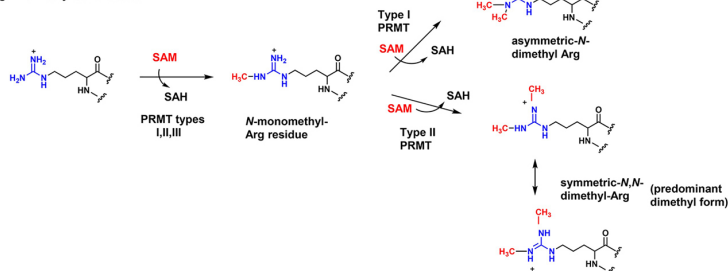
reactions of the guanidine group
in protein arginine residues



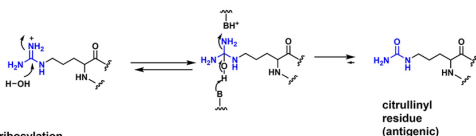
1. bacterial arginine kinase



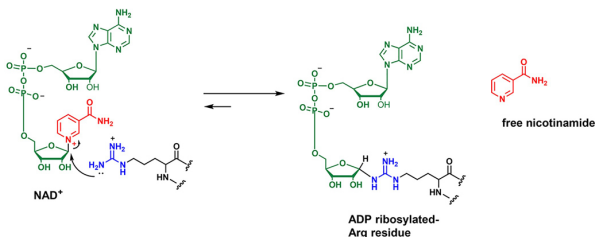
2. protein arginine methyl transferases



3. Peptidyl arginine deiminases



4. ADP-ribosylation



5. Nitroarginine

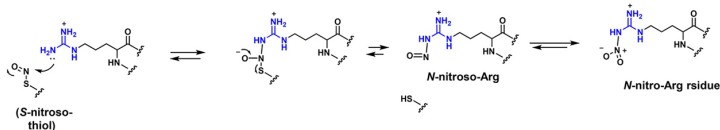


Figure 16.18 Five reactions of arginine guanidine or guanidinium side chains in proteins. (1) Arginine kinase-catalyzed N-phosphorylation by ATP. (2) Protein arginine N-methyltransferases (PRMTs) catalyse monomethylation and both symmetrical and unsymmetrical dimethylations on arginine guanidinium groups. (3) Protein arginine deiminases hydrolyze arginine to citrulline side chains that may act as neoantigens. (4) ADP ribosylation of guanidine groups. (5) Nitroarginine formation (may be nonenzymatic).

16.5.1 Bacterial Arginine Kinases

We noted in Chapter 8 that free arginines are among a handful of guanidinium metabolites in invertebrates that are enzymatically *N*-phosphorylated by ATP to create the N-PO_3^{2-} phosphoramidate linkages in product guanidinium phosphates such as *N*-phosphoarginine. These compounds, including *N*-phosphocreatine in mammalian muscle, are termed phosphagens because the kinase enzymes that form them also reverse readily with ADP to reform ATP. Thus, the *N*-phosphagens are thought to be metabolic high potential reservoirs of phosphoryl groups, a buffering source in tissues, such as skeletal muscle, with acute high demand for ATP expenditure.

It turns out that certain bacteria have cognate protein arginine kinases that phosphorylate arginine guanidinium side chains in specific client proteins. ⁴² These protein arginine kinases are homologous to the free arginine phosphagen kinases. The biological rationale for the posttranslational arginine residue *N*-phosphorylations appears to be that they serve as covalent marks that send the tagged proteins to chambered bacterial proteases, perhaps forerunners of eukaryotic proteasomes noted above in Section 16.4, for early proteolytic hydrolysis.

A comparative word about *N*-phosphorylated amino acid side chains in proteins: among histidine, lysine and arginine, one would expect all three, once formed, to retain phosphoryl group ($-\text{PO}_3^{2-}$) transfer potential to partner nucleophiles. *N*-Phosphohistidine residues are the main phosphoryl transfer metabolic currency in the bacterial two-component sensing systems noted in Section 16.3.2 above. *N*-Phosphoryl-lysine residues in proteins, on the other hand, are unexpectedly essentially absent from the inventory of protein phosphorylations, either as transient catalytic intermediates or stable posttranslational phosphoryl groups. Finally, the phosphoarginine residues are found as stable posttranslational modifications in a few bacterial proteins destined for early degradation, while free phosphoarginine in invertebrates is a buffering phosphagen. ⁴³

16.5.2 Arginine Guanidinium Methylations

Figure 16.11A, in addition to indicating sites of lysine residue N_6 methylations and acetylations among the four histone proteins that make up the core of nucleosomes in eukaryotic nuclei, also shows four arginines in histone H3 and one arginine in histone H4 that can be enzymatically methylated. The methyl donor is *S*-adenosylmethionine and one or two methyl groups can be delivered by distinct subsets of nine protein arginine *N*-methyltransferase (PRMT) enzymes (in single letter amino acid code Arg = R) (Figure 16.18 reaction pathway 2). ^{44,46} Of the nine, PRMT-1 carries approximately 80–85% of the flux of methyl groups to arginine residues. PRMT-1 is a tetramer as shown in Figure 16.19B, where the methyl donor *S*-adenosylmethionine is bound in two of the four subunits in a space filling view.

A protein arginine deiminase: citrulline residue formation via covalent thioimide intermediates

B tetrameric mouse PRMT1 with SAM bound to two subunits

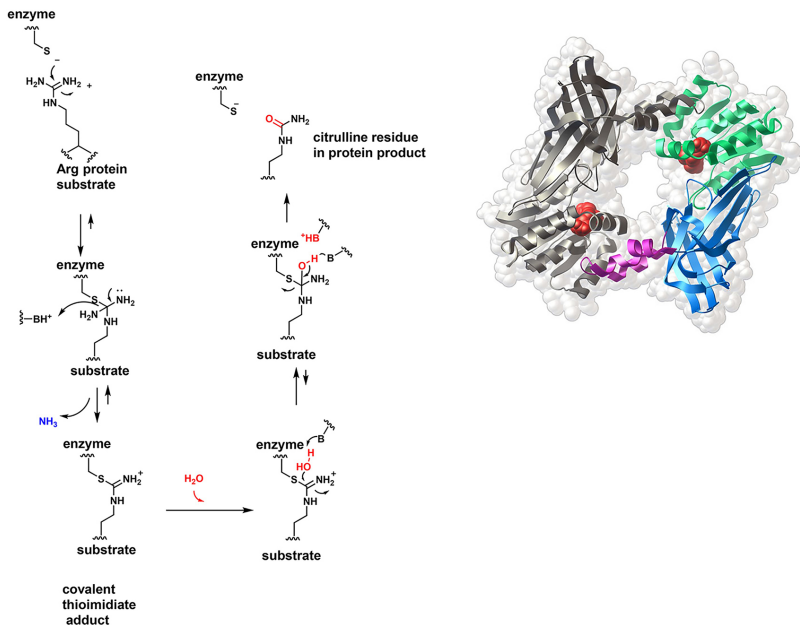


Figure 16.19 (A) The proposed reaction mechanism for protein arginine deiminases involves covalent thioimide adducts. (B) Mouse PRMT-1 tetramer with *S*-adenosylmethionine bound in two of the four subunits (PDB file 1OR8).

Three types of PRMTs are designated based on the composition and regiochemistry of methylation: *N*-monomethylation, or two types of dimethylation, symmetric or asymmetric. The asymmetric dimethyl arginine side chains have both methyl groups on the same nitrogen atom. The presumption from structural biology studies is that monomethyl-Arg residue side chains can fit into type I and Type II PRMTs with a different face of the methyl guanidine close to the SAM-methyl donor. A variety of PRMT inhibitors have been synthesized to evaluate whether specific PRMTs may be targets in diseases such as cancer.

16.5.3 Protein Arginine Deiminases

An arginine residue-specific posttranslational modification is enzymatic hydrolysis of the guanidinium group back to the biosynthetic precursor citrulline with a ureido group (Figure 16.18, reaction pathway 3). Citrulline is not one of the twenty proteinogenic amino acid building blocks and thus can be recognized as foreign and antigenic by the human immune system. The enzymes responsible for the deamination of arginine to citrulline residues are protein arginine deiminases. Hundreds of proteins are acted on by the deiminase family. Hyperactivity (termed hypercitrullination) is associated with a spectrum of autoimmune diseases.⁴⁷ Protein targets include histone H3 and H4, keratins, tubulins, vimentin and the abundant glycolytic enzyme enolase.

The mechanism of the four human peptidylarginine deiminases is thought to involve initial attack by a thiolate anion in the deiminase active site on the electrophilic guanidinium group of the arginine residue of a client protein. As depicted in Figure 16.19A, the resultant tetrahedral adduct can undergo proton transfers to eliminate

product NH_3 at the same time as a thioimide adduct links the deiminase enzyme to the arginine residue from the substrate protein. Catalyzed attack of water on that thioimide leads to another tetrahedral adduct from which elimination of the initially added enzyme thiolate releases the citrulline protein product. The chemistry is analogous to the amidino transferase involved in formation of guanidinoacetate described in Chapter 8.

16.5.4 Arginine ADP-ribosylations

A fourth type of arginine side chain posttranslational modification is enzyme-catalyzed ADP ribosylation from NAD^+ as the ADP donor. ⁴⁸ The substrate is the oxidized form of NAD^+ . It has a permanently quaternized pyridinium ion in the nicotinamide ring attached to C_1' of the ribose moiety of the NMN portion of NAD (Figure 16.18, reaction pathway 4), making the $\text{N}_1\text{--C}_1'$ pyridinium-ribose bond labile to capture by nucleophiles. A variety of ADP-ribosyl transferases initially isolated as bacterial protein toxins carry out these protein modifications on such host proteins as the P2X7 ion channel, the RAS GTPase and Arg₁₇₇ in actin, where the bulky substituent blocks actin polymerization. ⁴⁹

16.5.5 Nitroarginine Formation

One final type of arginine side chain modification that has been observed is formation of nitroarginine (reaction pathway 5 of Figure 16.18). In contrast to the previous four posttranslational modifications that are clearly enzyme-catalyzed, it may be that the nitration of arginine side chains is a secondary chemical consequence of nitric oxide formation, described in Chapter 14.

Among the indirect products of NOS upregulation are *S*-nitrosyl adducts, including *S*-nitrosyl glutathione. This glutathione adduct is diffusible through cells and a chemical donor of nitrosyl cations to nucleophiles. It may be that the source of nitroarginine residues in proteins at low fractional occupancy of sites is nitrosyl group transfer from *S*-nitrosyl glutathione. The resultant nitrosyl tyrosine may be the source of the eventual nitro group substituent. The nitroarginine residues are much less common than the corresponding nitrotyrosine residues noted in Chapter 14.

16.6 Asparagine, an Even More Unlikely Protein Side Chain Nucleophile

With its side chain nitrogen tied up as a resonance stabilized planar carboxamide group, the side chain of asparagine residues would seem, if anything, to be even less likely nucleophilic candidate than arginine side chains in both catalytic turnover mechanism for enzymes and in stable posttranslational modifications of proteins.

However, nature has chosen to use the carboxamido nitrogen of asparagine residues as the sole covalent linkage between proteins and *N*-glycan chains in hundreds to thousands of *N*-glycoproteins, predominantly in eukaryotic biology. ⁴ Some 2300 mouse proteins are *N*-glycosylated at an average of two asparagines per protein for 5000 known distinct sites. ⁵⁰ *N*-glycosylation happens in the endoplasmic reticulum and Golgi complex sections of cell secretory pathways. Because glycan chains are not templated in the way nucleic acids and proteins are encoded, there can be substantial variation of glycan chains, dependent on encounters during the secretory process with both early phase glycosyl transferases, remodeling glycosyl hydrolases, and maturation phase glycosyl transferases.

Yet, almost all *N*-glycoproteins follow a common early pathway utilizing a branched tetradecasaccharide, composition, Glc₃, Man₉, GlcNAc₂, where the two *N*-acetylglucosamines are most proximal to the asparagine side chain nitrogen (Figure 16.20). As in other topics, the *N*-glycoprotein literature is immense and even the rudimentary details of full processing of *N*-glycan chains are beyond the scope of this chapter.

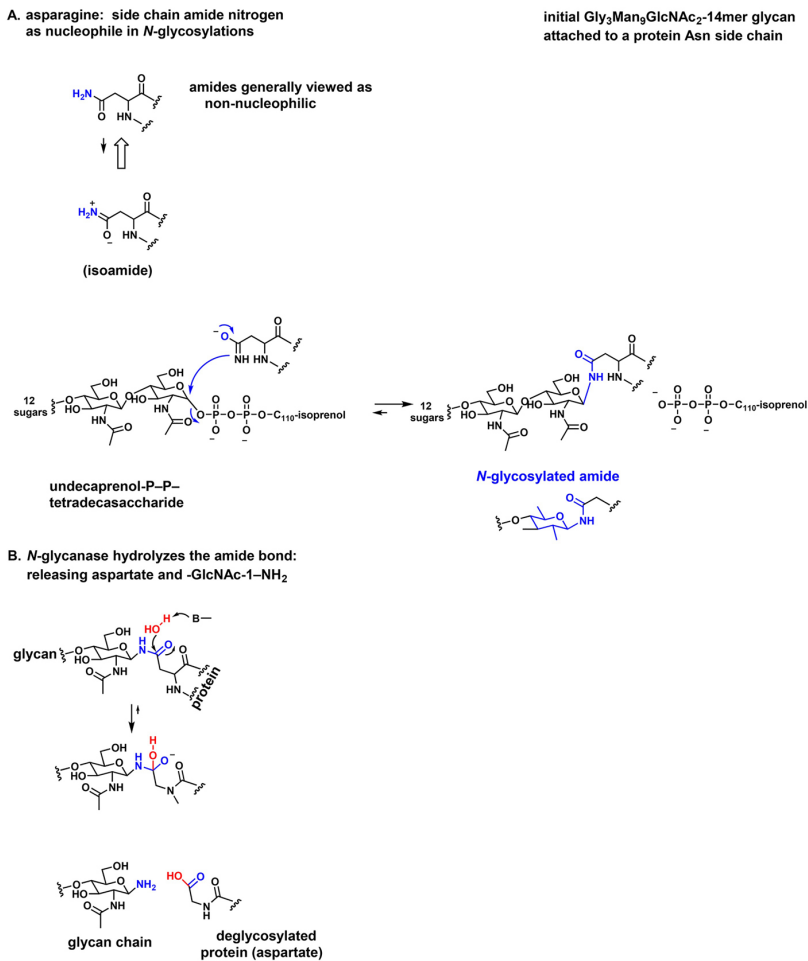


Figure 16.20 *N*-glycosylation of thousands of membrane and secreted proteins involve an asparagine carboxamide group, in a tripeptide sequence N-X-S/T to act as a nucleophile on a branched Glc₃Man₉GlcNAc₂-PP-dolichol as an electrophile. The Asn carboxamido nitrogen is an unlikely nitrogen nucleophile but may react as the minor isoamide anion in the active site of the membrane-embedded oligosaccharide transferase. The initial glycan chain is subsequently remodeled during passage of the *N*-glycosylated protein through the cell's secretory compartments but the Asn-*N*-β-GlcNAc linkage remains undisturbed. (B) In turn, when glycoproteins are taken up by cells and transit into lysosomes, the Asn-*N*-β-GlcNAc linkage is cleaved by the enzyme *N*-glycanase. The linkage is a glycosylated amide so *N*-glycanase functions as an amidase. This releases the protein with an aspartate rather than the starting asparagine side chain and yields

the glycan fragment with a 1-amino 2-*N*-acetyl glucose terminus. In the open chain form the 1-amino group can be readily eliminated as NH₃.

We focus only on the enzymatic step that attaches the initial tetradecasaccharyl chain to particular asparagine carboxamide nitrogen atoms in sequences summarized as N-X-S/T tripeptide regions. We also note the corresponding *N*-glycanase in the cleavage of that protein to a glycan N-C bond during glycoprotein degradation.

The cosubstrate tetradecasaccharide is preassembled in a pyrophosphate linkage to a long chain polyisoprenoid alcohol known as dolichol that is an endoplasmic reticulum (ER) membrane lipid. The tetradecasaccharyl-P-P-dolichol is oriented with the 14-sugar chain on the luminal side of the endoplasmic reticular membrane. Proteins that are *N*-glycosylated are modified cotranslationally before the protein substrates fold, as membrane-bound ribosomes extrude the nascent protein chain across the ER membrane. The oligosaccharyl transferase enzyme in humans is an eight-subunit, membrane associated complex that has been difficult to solubilize and study mechanistically.

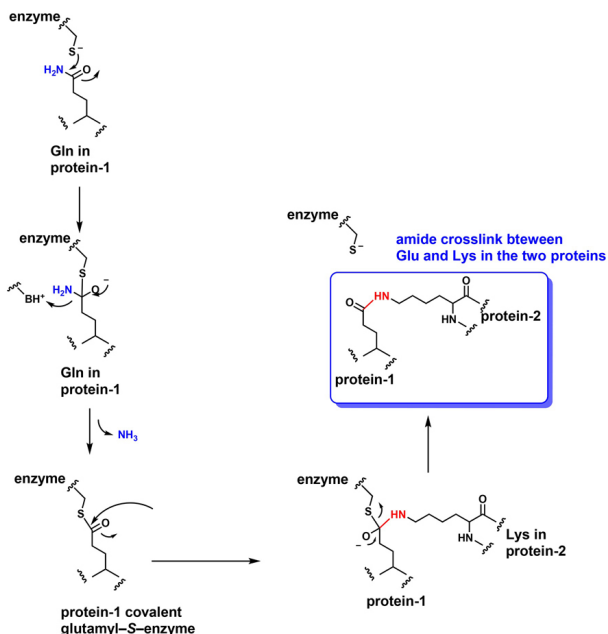
To obtain a nitrogen nucleophile on the relevant asparagine side chain it is possible that the oligosaccharyl transferase helps to populate the minor isoamide form of the protein substrate asparagine residue (Figure 16.20), perhaps by hydrogen bonding with the downstream Ser or Thr hydroxy side chain. A transition state with early dissociation of the GlcNAc₁-OPP-dolichol bond would create the C₁-oxocarbenium ion character that makes that C₁ carbon a functional electrophile in the reaction coordinate. Capture by the isoamide form of the side chain asparagine would result in a stable *N*-glycosylated amide. This is the universal covalent connection that links processed, mature glycan chains to asparagine residues in the many thousands of *N*-glycoproteins across many biological kingdoms.

At some later stage all those *N*-glycoproteins sitting in cell plasma membranes facing the exterior, or those that had been secreted extracellularly (e.g. antibodies), are taken back into cells through lysophagosomes and the sugar chain removed from the protein by the enzyme *N*-glycanase. The *N*-glycanase also removes *N*-glycan chains from glycoproteins that have misfolded and cannot reach compact native structures on the way out of the secretory pathway.

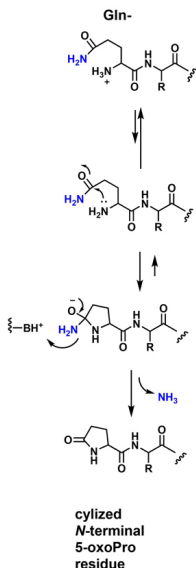
The *N*-glycanase 51 acts to hydrolyze the stable linkage between the protein asparagine nitrogen and C₁ of the proximal GlcNAc residues of a mature *N*-glycan chain (Figure 16.20B). From a chemical point of view this is a straightforward hydrolysis of an amide linkage. The initial products are the deglycosylated protein but with a side chain aspartate residue carboxylate not the starting amide of the asparagine residue. The initial amide nitrogen of asparagine has been left on the proximal GlcNAc₁ carbon 1 as 1-aminoGlcNAc-. That 1-aminoglucose scaffold is a labile hemiaminal that will ring open to the aldimine and lose ammonia as an observed secondary product.

A second example of posttranslational modification of the side chain amide of an asparagine residue occurs at Asn₄₁ of the conditional Rho GTPase, a member of a family of “small” GTPases involved in maintaining cell and organelle shape and dynamics. As shown in rection 3 of Figure 16.21, the ADP ribosyl moiety is transferred from NAD⁺ to the carboxamide nitrogen of Asn₄₁. 52 There are analogies to posttranslational ADP ribosylation of arginine residues noted in the previous section: net transfer of the electrophilic ADP ribosyl group to a weak nitrogen nucleophile.

1. transglutaminase crosslinking of proteins: Gln to Lys
by net transamidation



2. nonenzymatic cyclization of
N-terminal Gln to 5-OxoPro residue



3. posttranslational modifications of
Asn₄₁ and Gln₆₃ in Rho GTPase

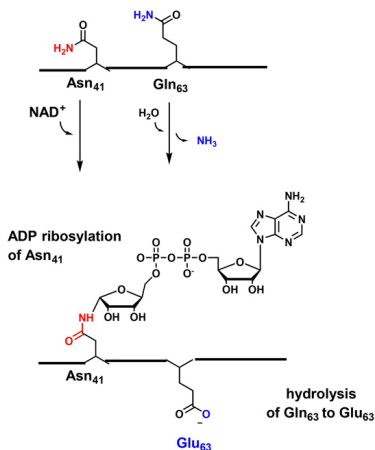


Figure 16.21 Three reactions of glutamine side chains in proteins. (1) Transglutaminase-mediated crosslinking of glutamine side chains to lysine side chains on a partner protein with release of NH₃ from the glutamine side occurs by capture of a glutamyl-S-enzyme by a lysine N₆-NH₂ on the second protein. (2)

Nonenzymatic cyclization of N-terminal glutamine residues to 5-oxoproline (gamma lactam) residues. (3). Posttranslational hydrolysis of Gln₆₃ in the Rho GTPase introduces a negative charge in the resultant glutamate₆₃ side chain and alters catalytic parameters.

16.7 Posttranslational Glutamine Deamidations

Like its congener asparagine side chain amide, glutamine residues are predicted to be largely inert as nitrogen nucleophiles, both as the free amino acid and when incorporated into protein backbones. In distinction from asparagine, we noted in Chapter 6, the widespread use of free glutamine's carboxamide group as a latent carrier of ammonia in a non-nucleophilic, nontoxic form. Glutaminase action *via* attack of an active site cysteine thiolate releases nascent NH₃ for amination reactions while forming a gamma glutamyl-S-enzyme intermediate which is then subjected to hydrolysis in every catalytic cycle.

This logic of release of ammonia *via* gamma glutamyl-S-enzyme formation is carried over to the transglutaminases, [53](#) posttranslational modification enzymes, that crosslink proteins covalently, e.g. in clotting biology. The cross links are isopeptide bonds between a glutamyl side chain on protein-1 and a lysine side chain on protein-2 ([Figure 16.21](#)). Inspection of starting protein-1 indicates the residue used as a donor for attack by a lysine N₆-NH₂ on protein-2 was actually a glutamine, not a glutamate. Furthermore, the crosslinking releases ammonia derived from the Gln carboxamide.

The transglutaminase name for these enzymes indicates common logic with glutaminases in the first half of a catalytic cycle. After binding to the enzyme, a protein-1 glutamine side chain amide is attacked by an active site cysteine thiolate anion in the transglutaminase active site. The tetrahedral adduct collapses in the forward direction to release NH₃ which exits as an observable product. The glutaminyl carbon skeleton is attached covalently to the attacking thiolate in the form of an acyl thioester, the gamma-glutamyl-S-enzyme. To this point the mechanism of transglutaminase and glutaminases (Chapter 6) are congruent.

In the second half reaction, deacylation of the acyl-S-enzyme intermediate, the mechanistic paths diverge. Where glutaminases utilize a water molecule as the deacylation nucleophile, transglutaminases bind a second protein substrate and orient a lysine N₆ amine side chain as the catalytically competent nucleophile. Acyl transfer of protein-1 to the lysine-N₆ of protein-2 creates the Glu-Lys covalent isopeptide crosslink.

One more comparison between glutaminases and transglutaminases. It is nascent NH₃ from glutaminase action that is the valuable product fragment for *in situ* aminations. In transglutaminases, the NH₃ is the waste product and it is the electrophilic glutamyl-S-enzyme that is the valuable intermediate for protein crosslinking.

A second net posttranslational hydrolysis of the glutamine side chain can occur in some proteins, including some antibodies, where glutamine is the N-terminal residue. As laid out in reaction 2 of [Figure 16.21](#), the free amino group of the N-terminal glutamine can attack the carbonyl group of the amide side chain intramolecularly. That tetrahedral adduct can expel NH₃ and result in formation of the cyclic amide that is 5-oxoproline. Lacking a free amino group at their N-terminus, the proteins that have undergone this cyclodeamination of Gln residues proved hard to sequence. This modification also makes them more hydrophobic. [54](#)

The third example, reaction 3, of [Figure 16.21](#) harks back to posttranslational regulatory modifications of the Rho GTPases. We noted in the preceding section that

Asn₄₁ can undergo enzymatic ADP ribosylation. The Rho enzyme can also undergo posttranslational deamidation (a glutaminase type reaction) on an internal glutamine, 55 Gln₆₃. That simple hydrolysis converts the neutral Gln side chain to the anionic side chain of glutamate which alters conformational dynamics within the folded enzyme structure. [Rho GTPase is a poster child for multiple PTMs on one single protein target. Four additional modifications are²: (1) O-glucosylation of Thr₃₇, (2) geranylgeranylation of Cys₁₉₀, (3) limited proteolysis at the C-terminus to liberate the prenylated Cys₁₉₀ as the new C-terminal residue, (4) O-methylation of the C-terminal carboxylate to the methyl ester to convert the carboxylate anion to neutral methyl ester, but none of these involve nitrogen chemical biology.]

16.8 Tryptophan C₂-Mannosylation as an Expression of Indole Ring Reactivity

A significant portion of Chapter 9 was given over to the reactivity of the indole ring of tryptophan, tryptophan dipeptides and tryptamine as a C₃ carbanion equivalent arising from the enamine part structure spanning N₁ and C₃ in the pyrrole ring moiety. Carbanion density at other positions in the indole ring can also be utilized as in the prenylation at C₄ that is the first committed step in the plant lysergic acid biosynthetic pathway.

Perhaps surprisingly, that enamine chemical potential is almost totally suppressed when tryptophan is incorporated into protein structures. There is one documented enzymatic reaction type where the indole ring of a tryptophan side chain in a protein acts as a carbon nucleophile when embedded in a W*XXW motif in mucins and extracellular metalloproteases of the A disintegrin and metalloproteases (ADAMs) family. The nucleophilic site on the W* residue is C₂ of the indole ring. The electrophile that it captures is C₁ of a mannosyl residue from UDP-mannose as glycosyl donor 5 (Figure 16.22). The function of this specific mono-mannosylation is unclear although it has been suggested that folding may be assisted.

Two kinds of Indole Side Chain Reactions

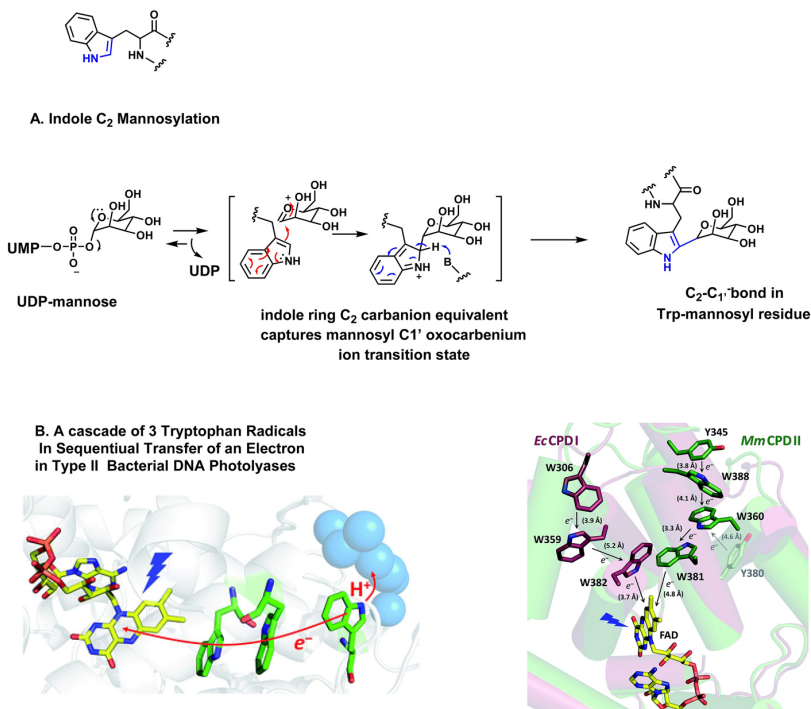


Figure 16.22 Two types of reactions of indole side chains in tryptophan residues in proteins: (A) C-Mannosylation of Trp at C₂ of the indole ring by a C₂ carbon nucleophile equivalent. Reproduced from [ref. 56](#) with permission from the Royal Society of Chemistry. (B) One-electron transfer via three tryptophan side chains in the light-driven DNA photolyase-driven cycloreversion of thymine dimer adducts in UV-damaged DNA. Reproduced from [ref. 56](#) with permission from the Royal Society of Chemistry.

Carbon C₂ of the indole ring is a site that can bear a partial negative charge in one of the resonance contributors to the indole ring structure, although it is typically much less reactive than C₃. A potential mechanism for the regiospecific C₂ mannosylation is shown in [Figure 16.22](#), hinting at the vast potential of the indole ring for carbanion chemical biology.

A second rarely observed property of tryptophan residues in proteins is to serve as transient radicals for electron flow over long distances within the protein matrix. One such role has been proposed for transfer of electrons from the nonheme iron center in ribonucleotide reductase to the active site for NDP reduction some 35 angstroms away in the partner subunit. Billions of the four nucleoside diphosphates have to be reduced to the 2' deoxyribonucleoside diphosphates in every cell cycle for DNA synthesis.

In the FAD-containing DNA photolyase enzyme that uses visible light as input as a substrate for radical reactions for reversion of thymine cyclobutane dimers in UV-damaged DNA, transient tryptophan indole ring radicals are detected. [56 Figure 16.22B](#) shows three tryptophans in both the *Escherichia coli* DNA lyase and the mouse DNA

lyase that channel or tunnel single electrons between the light absorbing flavin coenzyme as a photoantenna and the cyclobutene dimer to undergo one-electron cleavage mechanisms.

The use of light to cyclo-revert cyclobutane thymine dimers during DNA repair is spectacular chemistry. ⁵⁷ Along with radical based generation of DNA building blocks by ribonucleotide reductase, DNA photolyases reveal the biological appropriation of indole one-electron chemistry reaction manifolds as well as two-electron manifolds.

16.9 Overview

This survey of the catalytic and posttranslational chemistry of the six nitrogen side chains in the twenty amino acid residues found in proteins treats the covalent modifications, transient or stable, as a continuum of reactivity. This survey, in one sense, only scratches the surface of some 200 covalent posttranslational protein modifications. In another sense, it recapitulates the nitrogen chemical biology that plays out in imidazole, primary amine, guanidinium, carboxamide and indole low molecular weight metabolites.

Along with Ser-OH and Thr-OH side chains and the potent Cys thiolate nucleophile, these six amino acid nitrogen side chains carry most of the reaction flux where electrophilic group transfers occur in metabolism. To round out the catalytic inventory of nitrogen functional groups, the coenzyme forms of seven of the nitrogen heterocycles in B vitamins expand the scope of nitrogen-mediated chemical biology.

References

1. J. P. Klinman, A. Offenbacher and S. HuOrigins of enzyme catalysis: experimental findings for C-H activation, new models, and their relevance to prevailing theoretical constructs, *J. Am. Chem. Soc.*, 2017, **139**, 18409-18427.
2. C. T. Walsh, *Posttranslational Modification of Proteins: Expanding Nature's Inventory*, Roberts & Co, Englewood, Colorado, 2006.
3. M. A. C. Neves, M. Yaeger and R. AbagyanUnusual arginine formations in protein function and assembly: rings, strings and stacks, *J. Phys. Chem. B*, 2012, **116**, 7006-7013.
4. A. Larkin and B. ImperialiThe expanding horizons of asparagine-linked glycosylation, *Biochemistry*, 2011, **50**, 4411-4426.
5. Y. Ihara, Y. Inai, M. Ikezaki, I.-S. Li, S. Manabe and Y. Ito, C-mannosylation: modification on tryptophan in cellular proteins, in *Glycoscience: Biology and Medicine*, ed. N. Taniguchi, T. Endo, G. Hart, P. Seeberger and C. H. Wong, Springer, Tokyo, 2001.
6. J. S. BondProteases: history, discovery, and roles in health and disease, *J. Biol. Chem.*, 2019, **294**, 1643-1651.
7. K. Jung, L. Freid, S. Behr and R. HeermanHistidine kinases and response regulators in networks, *Curr. Opin. Microbiol.*, 2012, **15**, 118-124.
8. E. J. Capra and M. LaubThe evolution of two component signal transduction systems, *Annu. Rev. Microbiol.*, 2012, **66**, 325-347.
9. M. P. Bate, M. Goulian and W. F. DeGradoSignal transduction in histidine kinases: insights from new structures, *Structure*, 2015, **23**, 981-992.
10. C. T. Walsh, *The Chemical Biology of Phosphorus*, Royal Society of Chemistry Press, London UK, 2021.
11. S. Kwiatkowski and J. DrozakProtein histidine methylation, *Curr. Protein Pept. Sci.*, 2020, **21**, 675-689.

12. Y. Zheng, X. Zhang and H. LiMolecular basis for histidine N3 specific methylation of actin H73 by SETD3, *Cell Discovery*, 2020, **6**, 3.
13. Q. Al-Hadid, K. Roy, W. Munroe, M. C. Dzialo, G. F. Chanfreau and S. G. ClarkeHistidine methylation of yeast protein Rpl3p is required for proper 60S subunit assembly, *Mol. Cell. Biol.*, 2014, **34**, 2903-2916.
14. R. A. JohnPyridoxal-P dependent enzymes, *Biochim. Biophys. Acta*, 1995, **1248**, 81-96.
15. P. K. Mehta and P. ChristenThe molecular evolution of pyridoxal-phosphate-dependent enzymes, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 1999, **74**, 129-184.
16. M. C. Ho, J. Menetret, H. Tsuruta and K. N. AllenThe origin of the electrostatic perturbation in acetoacetate decarboxylase, *Nat. Rev. Microbiol.*, 2009, **459**, 393-397.
17. J. J. Marsh and H. LeebherzFructose-bis-phosphate aldolases: an evolutionary history, *Trends Biochem. Sci.*, 1992, **17**, 110-113.
18. S. ShulmanDNA ligases: progress and prospects, *J. Biol. Chem.*, 2009, **284**, 17365-17369.
19. S. M. Carlson and O. GorzaniEmerging technologies to map the protein methylome, *J. Mol. Biol.*, 2014, **426**, 3350-3362.
20. R. C. Trievel, E. Flynn, R. L. Houtz and J. M. HurleyMechanism of multiple lysine methylation by the SET domain enzyme Rubsico LSMT, *Nat. Struct. Mol. Biol.*, 2003, **10**, 545-552.
21. S. G. ClarkeProtein methylation at the surface and buried deep: thinking outside the histone box, *Trends Biochem. Sci.*, 2013, **38**, 243-252.
22. B. Keppler and T. ArcherChromatin-modifying enzymes as therapeutic targets-part I, *Expert Opin. Ther. Targets*, 2008, **12**, 1301-1312.
23. B. J. G. E. Pieters, M. Wuts, J. Poater, K. Kumar, P. B. White, J. J. A. G. Kamps, W. Sherman, G. J. M. Pruijn, R. S. Paton, T. Beuming, F. M. Bickelhaupt and J. MecinovicMechanism of bimolecular recognition of trimethyllysine by the fluorinated aromatic cage of KDM5A PHD3 finger, *Commun. Chem.*, 2020, **3**, 69.
24. B. M. Miller and A. GulickStructural biology of non-ribosomal peptide synthetases, *Methods Mol. Biol.*, 2017, **1401**, 3-29.
25. S. Mori, A. Pang, T. A. Lundy, A. Garzan, O. V. Tsodikov and S. Garneau-TsodikovaStructural basis for backbone N-methylation by an interrupted adenylation domain, *Nat. Chem. Biol.*, 2018, **14**, 428-430.
26. H. Song and J. NaismithEnzymatic methylation of the amide bond, *Curr. Opin. Struct. Biol.*, 2020, **65**, 79-99.
27. J. Rice, The Histone Code, www.histonecode.com 2020.
28. T. Jenuwein and D. AllisTranslating the histone code, *Science*, 2001, **293**, 1074-1080.
29. G. Molazzo, D. Mrecatelli, G. Di Muzio, L. Triboli, P. De Rosa, G. Perini and F. M. GorgiHistone deacetylases (HDACs): evolution, specificity, role in transcriptional complexes, and pharmacology actionability, *Genes*, 2020, **11**, 5560604.
30. Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, Y. Casero and Y. ShiHistone demethylation mediated by the nuclear amine oxidase homolog LSD1, *Cell*, 2004, **119**, 941-953.
31. Y.-I. Tsukada, J. Fang, H. Erdumant-Bromages, M. E. Warren, C. H. Borchers, P. Tempst and Y. ZhangHistone demethylation by a family of JmjC domain-containing proteins, *Nat. Rev. Microbiol.*, 2006, **439**, 811-816.
32. T. FaraziThe biology and enzymology of protein N-myristoylation, *J. Biol. Chem.*, 2001, **276**, 39501-39504.
33. K. D. WilkinsonThe discovery of ubiquitin-dependent proteolysis, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15280-15282.
34. C. M. PickartMechanisms underlying ubiquitylation, *Annu. Rev. Biochem.*, 2001,

35. R. T. HaySUMO; a history of modification, *Mol. Cell*, 2005, **18**, 1-12.
36. M. J. Matunis, E. Coutavits and G. BlobelA novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase activating protein RanGAP-1 between the cytosol and the nuclear pore complex, *J. Cell Biol.*, 1996, **135**, 1457-1470.
37. K. Wollenberg and J. SwaffieldEvolution of proteasomal ATPases, *Mol. Biol. Evol.*, 2001, **18**, 962-974.
38. R. T. Sauer and T. BakerAAA+ ATPases: ATP-fueled machines of protein destruction, *Annu. Rev. Biochem.*, 2011, **80**, 587-612.
39. T. E. T. Mevissen and D. KomanderMechanisms of deubiquitinase specificity and regulation, *Annu. Rev. Biochem.*, 2017, **86**, 159-192.
40. A. Nayak and S. MullerSUMO-specific proteases/isopeptidases: SENPs and beyond, *Genome Biol.*, 2014, **15**, 422.
41. D. J. Slade, V. Subramanian, J. Fuhrmann and P. R. ThomsonChemical and biological methods to detect posttranslational modifications of arginine, *Biopolymers*, 2014, **101**, 133-143.
42. M. J. Suskiewicz, B. Hajdusits, R. Beveridge, A. Heuck, L. Da Vu, R. Kurzbauer, K. Kauer, V. Thoeny, K. Rumpel, K. Mechtler, A. Meinhart and T. ClausenStructure of McsB, a protein kinase for regulated arginine phosphorylation, *Nat. Chem. Biol.*, 2019, **15**, 510-518.
43. A. Azzi, S. Clark, W. R. Ellington and M. S. ChapmanThe role of phosphagen specificity loops in arginine kinase, *Protein Sci.*, 2009, **13**, 575-585.
44. L. Sun, M. Wang, Z. Lv, N. Yang, Y. Liu, S. Bao, W. Gong and R.-M. XuStructural insights into protein arginine symmetric dimethylation by PRMT5, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 20538-20543.
45. M. D. Fulton, T. Brown and Y. G. ZhengThe biological axis of protein arginine methylation and asymmetric dimethylarginine, *Int. J. Mol. Sci.*, 2019, **20**, 3322.
46. S. Tewary, Y. Zheng and M.-C. HoProtein arginine methyltransferase: insights into the enzyme structure and mechanism at the atomic level, *Cell. Mol. Life Sci.*, 2019, **76**, 2917-2932.
47. S. Mondal and P. ThompsonProtein arginine deiminases (PADs): biochemistry and chemical biology of protein citrullination, *Acc. Chem. Res.*, 2019, **52**, 818-832.
48. S. Laing, M. Unger, F. Koch-Nolte and F. HaagADP-ribosylation of arginine, *Amino Acids*, 2011, **41**, 257-269.
49. Cytoskeleton Inc, *Posttranslational Modifications of Actin*, 2020, www.cytoskeleton.com.
50. M. AebiN-linked protein glycosylation in the ER, *Biochim. Biophys. Acta*, 2013, **1833**, 2430-2437.
51. T. SuzukiThe cytoplasmic peptide:N-glycanase (Ngly1)-basic science encounters a human genetic disorder, *J. Biochem.*, 2015, **157**, 23-34.
52. H. Genth, R. Gerhard, A. Maeda, M. Amano, K. Kaibuchi, K. Aktories and I. JustEntrapment of Rho ADP-ribosylated by Clostridium botulinum C3 exoenzyme in the Rho-guanine nucleotide dissociation-inhibitor-1 complex, *J. Biol. Chem.*, 2003, **278**, 28523-28527.
53. M. Griffin, R. Casadio and C. M. BergaminiTransglutaminases: nature's biological glues, *Biochem. J.*, 2002, **368**(pt 2), 377-396.
54. D. Galante, F. Ruggeri, G. Dietkler, F. Pellistri, A. Gatta, T. Florio, A. Perico and V. D-ArrigoA critical concentration of N-terminal pyroglutamylated amyloid beta drives the misfolding of Ab1-42 into more toxic aggregates, *Int. J. Biochem. Cell Biol.*, 2016, **79**, 261-270.
55. G. Schmidt, P. Sehr, M. Wilm, J. Selzer, M. Mann and K. AktoriesGln63 of Rho is deamidated by E.coli cytotoxic necrotizing factor-1, *Nature*, 1997, **387**, 725-729.

56. P. Muller, E. Ignatz, S. Kiontke, K. Brettel and L.-O. Essen Sub-nanosecond tryptophan radical deprotonation mediated by a protein-bound water cluster in class II DNA photolyases, *Chem. Sci.*, 2018, **9**, 1200-1212.
57. A. Sancar It is chemistry but not your grandfather's chemistry, *Biochemistry*, 2017, **56**, 1-2.
58. Z. Xiaolin *et al.*, Histone Lysine-Specific Methyltransferases and Demethylases in Carcinogenesis: New Targets for Cancer Therapy and Prevention, *Curr. Cancer Drug Targets*, 2013, **13**, 558-579.
59. S. Martinet *et al.*, Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction, *Nat. Rev. Neurosci.*, 2007, **8**, 948-959.
60. T. Hagai and Y. Levy Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 2002.

Subject Index

- acetoacetate decarboxylase, 358
- acylations, 156–158, 356, 367–373
- acyl CoA metabolites, 365–367
- adenosine monophosphate (AMP), 179–182
- adenosine triphosphate (ATP), 113, 315
- ADP-forming amide bond synthetases, 325–327
- alkylations, 356
- amides, 174–176, 182–184
- amidines, 139–141, 174–176
- amidino transfer, 150, 151
- amines, 153–155
 - acylations, 156–158
 - electrophilic group transfer reactions, 155–156
 - glycosylations, 162–163
 - metabolic amine alkylations, 158–162
 - to *N*-hydroxylamines, 246–249
 - N*-hydroxylations, 165–166
 - to nitroso and nitro groups, 252–256
 - N*-sulfurylation and *N*-phosphorylation, 163–165
 - six-electron oxidations of, 252–256
 - substrate amino group capture, 155–156
 - two-electron oxygenation of, 246–249
- 1-amino-1-carboxy-cyclopropane, 283–284
- amino acid dehydrogenases, 129–130
- amino acid monomer activation, 320–321
 - ammonium cations to amine nucleophiles, deprotonation, 321–322
 - carboxylate activation, 322
- amino acid oxidases, 129–130
- amino acid transformations, 205–207
- aminoacyl thioester strategy, 344
- aminoadipyl-cysteiny-D-valine (ACV) synthetase
 - elongation module, 340–341
 - initiation module, 340
 - termination module, 341
 - three module NRPS assembly line, 340
- 1-aminocyclopropane-1-carboxylate (ACC), 284
- amino groups, organic scaffolds, 121

transaminase logic, [121–126](#)

aminolevulinate (ALA), [221](#)

ammonia (NH₃), [1](#), [2](#), [26](#), [27](#), [28](#)

- crossover metabolite, [134](#)
- enzymatic capture of, [14–16](#)
- steady state nontoxic level, [137–139](#)

ammonia oxidase, [52–55](#)

ammonia oxidizing bacteria, [49](#), [50](#)

ammonification, [69](#)

AMPylations, [356](#)

- DNA ligase catalytic repair, [358–361](#)

anaerobic bacterial phylum plantomycetes, [62–63](#)

anammox pathway

- anaerobic bacterial phylum plantomycetes, [62–63](#)
- hydrazine dehydrogenase, [66–68](#)
- hydrazine synthase, [63–66](#)
- logic, [61](#)

ardeemin, [192](#)

arginase, urea cycle, [148–150](#)

arginine

- arginine ADP-ribosylations, [377–379](#)
- arginine guanidinium methylations, [375–377](#)
- bacterial arginine kinases, [375](#)
- nitroarginine formation, [379](#)
- protein arginine deiminases, [377](#)
- protein side chain nucleophile, [373–375](#)

arginine ADP-ribosylations, [377–379](#)

arginine guanidinium methylations, [375–377](#)

arginine *N*-hydroxylation, [260–263](#)

asparagine, protein side chain nucleophile, [379–382](#)

aspartate C₂-amino group, [146–148](#)

aspartate transcarbamylase, [144–146](#)

assimilatory nitrate reduction, [29](#)

atomic nitrogen, [6–8](#)

ATP grasp family, [325–327](#)

auto-posttranslational modification process, [226](#)

azote, [6](#)

bacterial arginine kinases, [375](#)

bacterial nitrogenases, [31–32](#)

base–acid catalysis

- cysteine proteases, [351–353](#)
- serine proteases, [351–353](#)

biogeochemical cycle diagrams, [92](#)

biotin (vitamin B₇), [207–210](#)

- calicheamicin, 249, 250
- carbamate, 141–142
- carbamic acid, 142–144
- carbamyl phosphate, 141–142, 142–144
- carbamyl transfers, 141–142
- carbanion–enamine equivalent, 185
- carbinolamine, 114
- carbon dioxide, 283–284
- cationic thiazolium ring, 198
- 7-chlorotryptophan, 192
- citrulline, 146–148, 377
- copper, 89
- cyanide ion, 271–273, 283–284
- cycloalanine nitrile hydrolases, 279
- cycloalanine synthase, 277–279
- cyanogenesis strategies, 283
 - bacterial amino acid fragmentations, 284
 - glycine-fragmenting HCN synthase, 284–286
 - tyrosine, enzymatic fragmentation, 286–287
 - enzymatic oxidative fragmentation, 283–284
- cyanogenic glycosides, 276–277
- cyclic amines, 182–184
- cyclic metabolites, 183
- cytochrome P450 monooxygenases, 312
- D-cycloserine, 260–263
- 7-deazaguanine acid, 282
- 7-deazaguanine scaffolds, 290
- denitrification, 6, 26, 29–31
 - denitrifying nitrate reductase, 81
 - denitrifying nitrite reductases, 81–83
 - metallocenter inventory, 86–90
 - net ten-electron reduction, 81
 - nitric oxide reductase, denitrifying bacteria, 83–86
 - nitrous oxide reductase, 86
 - one-electron reductions, 81–83
- dihydroheteroaromatic ring system, 204
- dihydronicotinamides, 202–205
- dinitrogen, biological fixation, 31
 - alternative to nitrogenase, 32
 - Haber–Bosch industrial process, 32
- dinitrogen gas, 6, 25, 27
- dissolved nitrate anions, 6, 27
- DNA, nitrogen heterocyclic bases, 3
- DNA base pairs, 174–176
- eight-electron nitrogen redox traverse, 12–13
- electrophilic group transfer reactions, 155–156

enamine intermediates, [170](#)

enamines, [153–155](#)

enediynes, antitumor antibiotic calicheamicin, [249](#)

enzymatic amination, [109–113](#)

 ammonia organification, amide linkage, [114–116](#)

 glutamate dehydrogenase, [113–114](#)

 glutamine oxoglutarate aminotransferase, [116](#)

 glutamine synthetase, [114–116](#)

 imines and amides, [109–113](#)

 reductive amination, substrate carbonyl groups, [113–114](#)

enzymatic ornithine hydroxylation, [249](#)

ethylene, [283–284](#)

flavin adenine dinucleotide (FAD), [201](#)

flavin (B₂) coenzymes

 nature's step-down redox transformer, [201–202](#)

 tricyclic ring system of, [201–202](#)

flavin-dependent monooxygenase, [252](#)

flavin mononucleotide (FMN), [201](#)

flavoenzyme oxygenases, [165](#)

folate (vitamin B₉), [210–212](#)

functional groups, metabolism, [16](#)

gaseous ammonia, [6](#)

glucosinolates, [258–259](#)

glutamate dehydrogenase, [105](#), [113–114](#)

glutamate synthase, [105](#)

glutamine, latent ammonia carrier

 asparagine and aspartate vs. glutamate and glutamine, [119–121](#)

 enzymatic aminations, general strategy, [116–119](#)

glutamine oxoglutarate aminotransferase (GOGAT), [116](#)

glutamine synthase, [105](#)

glutamine synthetase, [114–116](#)

glycosylations, [162–163](#)

guanidines, [139–141](#), [174–176](#)

guanosine monophosphate (GMP), [179–182](#)

Haber–Bosch industrial process, [32](#)

heme biosynthesis, [221](#)

 aminolevulinate

 to porphobilinogen, [223–225](#)

 routes to, [221–223](#)

 chelates, [234–236](#)

 imine and enamine chemistry, [223–225](#)

 porphobilinogen, [226](#)

 to hydroxybilane, [226–230](#)

 uroporphyrinogen-III synthase, [230–231](#)

heme cofactor, 219

hemeprotein cytochrome P450 monooxygenases, 300

histidine

- as acid–base and nucleophilic catalyst, 351

- autokinases, 353–355

- imidazole ring, posttranslational methylation, 355

- as nucleophile, 353–355

human B vitamins, 195–198

human flavin monooxygenase (FMO) isozymes, 267

hydrazine dehydrogenase, 66–68

hydrazine synthase, 63–66

hydrogen cyanide (hydrocyanic acid), 271–273

hydroxylase synthase, 133

hydroxybilane synthase, 226

hydroxylamine oxidoreductase, 55–58

imidazole, 316

imine–enamine tautomerizations, 154, 356

imine–iminium ion reactions, 166–167

- four-iminium intermediates, 167–168

- ornithine to cocaine, 167–168

- tyrosine to tetrahydroisoquinoline, 168–169

imines, 153–155

iminium ion formations, 356–358

iminoglutarate reductase, 113

indole prenylations, regiochemistry, 191–192

indole ring reactivity, 384–386

information-rich ubiquitin, 367–373

inorganic metabolites, 3, 4

inorganic nitrogen atoms, 292–293

inorganic nitrogen chemical biology, 28

inorganic nitrogen metabolism, 12–13, 92

inorganic nitrogen redox pathway enzymes, 100

interconnected functional groups, 17

iron-based vs. flavin-based *N*-oxygenases, 244–246

isonitriles, biological formation, 288–289

isopeptide bond, 326

iterative methylations, 361–363

- peptide bond nitrogens, *N*-methylation, 363–365

lithotrophic energy metabolism, 5

lysine, 249, 251, 356

- to hydrazine derivatives, 260

metallated tetrapyrrole macrocycles, 218–219

methionine synthase, 212, 215

methylene imidazolone (MIO) cofactor, 128

microbial nitrification, 49

mixed acyl–phosphoric anhydrides, 322

thermodynamic activation, 322–324

molecular nitrogen, 6–8

molybdenum, 58, 59, 102

molybdenum cofactors, 103

molybdenum nitrogenase, 35

net amidation reaction, 15

N-hydroxy-pipecolate, 249

N-hydroxyurea, 263

nicotinic acid mononucleotide (NAMN), 202

nitrate ions, biological reduction, 72–74

nitrate reductases, 74–77

nitrate reductions, 29, 30

assimilation and eight-electron reduction, 74

nitrate reductases, 74–77

nitrite reductases, 77–80

nitric acid, 28

nitric oxide biology, 306–307

direct and indirect effects, eukaryotes, 307–310

guanylate cyclase, interaction, 311

nitric oxide oxidase, 52, 58

nitric oxide reductase

denitrifying bacteria, 83–86

nitric oxide synthase

arginine guanidium group, 263–265

coenzymes and P450 type heme, 265–266

three-electron *N*-oxidation, 263–265

nitrification, 29–31

ammonia oxidation to nitrate, 48–52

ammonia to nitrate oxidative pathway, 60

nitrifiers, 26

nitrite oxidase, 58–60

nitrite oxidizing bacteria, 58–60

nitrite reductases, 77–80

six-electron reduction, 78–80

nitroarginine formation, 379

nitrogenase enzyme, 1, 25

nitrogenase mechanisms, 41–45

hydrazine and diazine species, 41

nitrogen swinging, M cluster, 43–45

storage of electrons, M cluster, 41–43

nitrogenase redox cofactor inventories, 37–40

iron–sulfur clusters, 37–38

P, M and V cluster, nitrogenases, 38–40

nitrogen biogeochemical cycling 1, 10–12

nitrogen biogeochemical cycling 2, 13–14

nitrogen-centric chemical biology, 22

nitrogen-fixing bacteria, 33

nitrogenase mechanisms, 41–45

nitrogenase redox cofactor inventories, 37–40

nitrogenase structures, 33–37

reaction stoichiometries, 33–37

nitrogen heterocycles, 2, 172–174, 216

AMP and GMP, 179–182

central structural and catalytic groups, 16–19

DNA base pairs and, 174–176

in metabolism, 173

nitrogens origin, parsing, 176–182

pyrimidines and purines, 176–182

uridine and cytidine, 176–179

nitrogen hydride biochemistry, 45

nitrogen inorganics, 94–97

nitrogen oxidation states, 8–10

nitrogen oxide radical, 19

nitrogen oxide vs. nitrogen hydride chemical biology, 97–99

nitrogen proteomics, 21–22

catalysis vs. posttranslational modifications, 21–22

nitrogen redox enzymology, 19–21

one-electron vs. two-electron, 19–21

nitrogen redox states, 94

nitrones, 256–257

nitrous oxide reductase, 86

nonribosomal peptide synthetase assembly lines, 330–337

architecture, 337–339

nonribosomal peptide variants, 341–343

nucleophilic reactivity, 8

obligate hydride transfer agents, 202–205

odd electron oxidations, 20

omphalotin A, 364

one-electron chemistry, 20

one-electron transfers, 292–293

orbital filling diagram, 7

organic nitrogen chemical biology, 14–16, 102

organic nitrogen radical metabolites, 296

- carbon and nitrene transfer, cytochrome P450s, [301–303](#)
- cyanogenesis, [303–304](#)
- electron flow to FADH₂, [297](#)
- FADH₂ dependent oxygenation and transfer, [297–299](#)
- nitric oxide synthase, one-electron transfer mechanisms, [304–306](#)
- nitrogen radical coupling, [301–303](#)
- nitrogen radicals, cyanide ion formation, [303–304](#)
- one-electron transfer, cytochrome P450 N-oxygenations, [299–301](#)

organic nitrogen scaffolds

- PLP-dependent transaminases, redox catalysts, [295](#)
- two-electron vs. one-electron redox manifolds, [293–294](#)

organification, [113](#)

- ornithine, to hydrazine derivatives, [260](#)

- oxidative pathway, [94–97](#)

- oximes, [256–257](#)

- peptide bonds, [319–320](#)

- amino acid monomer activation, [320–321](#)

- chain formation, ribosomal initiation, [327–328](#)

- initiation, elongation and termination, ribosomal RNA, [327](#)

- peptidyl chain elongation, ribosome, [328–330](#)

- peptidyl chain termination, ribosome, [330](#)

- phlogisticated air, [6](#)

- phosphagen, [150](#)

- physostigmine biosynthetic strategy, [185–186](#)

- pigments of life, [2](#), [193](#), [220](#)

- plant cytochrome P450

- amino acids to cyanohydrin glycosides

- aldoxime to cyanohydrin acetals, [275–276](#)

- subclass of Cyp79 heme N-oxygenases, [274–275](#)

- porphobilinogen, [2](#)

- porphobilinogen synthase, [225](#)

- posttranslational glutamine deamidations, [382–384](#)

- posttranslational modifications (PTMs), [314](#)

- precorrin [2](#), [239](#)

- protein arginine deiminases, [377](#)

- proteinogenic amino acid building blocks, [347–350](#)

- protein residue nitrogen side chains, [350–351](#)

- pyrazolylalanine, [184](#)

- pyridoxal-phosphate (PLP) (B₆ coenzyme), [121](#), [205–207](#), [206](#), [247](#), [295](#), [356](#)

- pyrrole ring reactivity, [227](#)

- quaternized thiazolium ring, [198–200](#)

- reactions that release ammonia

- amidino group deaminases, [130–133](#)

- amino acid deaminases, [126–129](#)

- amino acid dehydrogenases and oxidases, deamination, [129–130](#)

- ammonia lyases, [126–129](#)
- purines, pyrimidines and nucleic acids, [130–133](#)
- tetrapyrrole macrocycle biosynthesis, [133–134](#)
- redox active cofactors, [51](#)
- redox cofactor inventories, [101–102](#)
- reductive amination, [15](#)
- reductive assimilation, [29–31](#)
- reductive metabolic pathways, [94–97](#)
- reversible posttranslational acylations, [365–367](#)
- riboflavin, [201](#), [202](#)
- siderophores, [246](#), [247](#)
- siroheme, [236–239](#)
- small ubiquitin like molecule (SUMO) isoforms, [367–373](#)
- symbiotic nitrogen fixing bacteria, [34](#)
- tetrahydro- β -carboline, [188](#)
- tetrapyrrole macrocycles, [106](#)
 - biosynthesis, [133–134](#)
 - biosynthetic branchpoints, protoporphyrin IX to chlorophylls, [239–241](#)
 - biosynthetic branchpoints, uroporphyrinogen-III to siroheme, [236–239](#)
- thiamin-PP, [198–200](#)
- three-electron oxidation, [100](#)
- toyocamycin, [280–283](#)
 - nitrile formation and reduction, [280–283](#)
- transaminase logic, [121–126](#)
 - second half reaction, [125–126](#)
 - two half reactions, [125](#)
- transamination, [122](#)
- tricyclic isoalloxazine ring, [196](#), [201](#), [216](#)
- tryptamine
 - acylation, [185](#)
 - alkylation, [185](#)
 - imine to beta carboline scaffold, [186–188](#)
- tryptophan, [192](#)
 - diketopiperazine formations, [188–191](#)
 - internal cyclizations, [188–191](#)
- tryptophan C₂-mannosylation, [384–386](#)
- two-step cyanogenesis, [276–277](#)
- ureas, [139–141](#)
- vitamin B₁, [198–200](#)
- vitamin B₃
 - coenzyme forms of, [202–205](#)
 - dihydronicotinamides, [202–205](#)
 - obligate hydride transfer agents, [202–205](#)

vitamin B₁₂

modified tetrapyrrole macrocycle, [212–215](#)

organometallic coenzyme forms, [212–215](#)

Watson–Crick scheme, [174](#)